Five Empirical Essays on Competition Policy and Health Economics

Doctoral Thesis

submitted to Justus-Liebig-University Giessen,

Department of Economics

1. supervisor: Prof. Dr. Georg Götz

2. supervisor: Prof. Dr. Peter Winker

Phil-Adrian Klotz*

February 14, 2024

*Phil-Adrian Klotz is a research associate at the Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen, Germany. E-mail: phil.a.klotz@wirtschaft.uni-giessen.de. Tel.: +49-641 99 22060.

Acknowledgements

My thanks go to all the people who supported me during the last five years. All the essays of this doctoral thesis have been widely discussed and I am deeply thankful to all those who gave me input to my work at conferences and workshops. I am especially grateful to my supervisor Georg Götz for his guidance, support, and invaluable insights throughout my research process. Our extensive and productive discussions, coupled with his thorough feedback and suggestions on earlier versions of the various papers, offered valuable input and numerous fresh perspectives. I am also thankful to my second supervisor Prof. Winker for his constructive feedback and thoughtful suggestions on earlier versions of some of my essays. I extend my appreciation to my co-authors Chiara Drolsbach, Maximilian Gail, Daniel Herold, and Jan Schäfer who collaborated with me on various aspects of this research. Working and discussing with them was always fruitful and stimulating. Moreover, I owe my colleagues at the Department for Economics of the Justus-Liebig-University Gießen and especially at the Chair of Industrial Organization, Regulation and Antitrust sincere gratitude. The whole team provided a very stimulating research environment for my dissertation project in the last five years. Last but not least, I would like to thank my family and my friends for the support and encouragement throughout the last years.

Table of Contents

1 Introduction

This doctoral thesis and its associated papers address empirical research questions in the fields of competition policy and health economics. In all five papers, empirical microeconomic tools are applied to identify and measure causal links. Mostly, (quasi) natural experiments are employed to estimate the impact of policy interventions on market outcomes.

The connection between the five papers in this thesis is that causal inference methods are used to analyze economic policy issues. Causal inference is the process of uncovering causal effects by estimating the impact of events and choices on a given outcome of interest (Cunningham, 2021). In the papers of this thesis, observational data is used to answer the individual research questions. However, correlations in this type of data are mostly not reflecting causal relationships because the variables are based on choices of individuals which create spurious correlations with other things (Huntington-Klein, 2021). Hence, causal inference methods are needed to identify causal links based on certain assumptions.

The first three papers of this doctoral thesis empirically analyze the behavior of firms and consumers in different markets. The first paper *The Substitutability Between Brick-and-Mortar Stores and e-Commerce - The Case of Books* studies the competition between online and offline retailers using the German book market as an example. This market is particularly well-suited for such an analysis as it is characterized by fixed book prices such that the price-dimension plays no role in the comparison between the two channels. The present paper analyzes how closures of physical bookstores affect book sales in order to study to what extent consumers perceive the online and offline channels as substitutes by applying an instrumental variable approach. The results of this paper imply that there is a imperfect substitutability between the two sales channels, while the magnitude of the effect depends on the individual book genres.

The second paper *The Impact of the Agency Model on E-book Prices: Evidence from the UK* also addresses the book market. This study uses an unique data set of e-Book prices to empirically analyze the effect of the widely used agency model on retail prices. Using agency arrangement means that suppliers pay retailers sales royalties to distribute products at prices determined by suppliers (De los Santos & Wildenbeest, 2017). This type of vertical contract is especially prevalent in online markets and discussed in both, academia and practice (e.g., Foros et al., 2017; Johnson, 2017; Li and Moul, 2015). For instance, the Department of Justice (DOJ) sued Apple and five publishing houses for conspiring to raise e-Book prices by using the agency model in conjunction with most-favored nation (MFN) clauses in 2012.[1] The research goal in the present paper is to analyze the price effect of the agency model by comparing e-Books sold under this pricing arrangement with e-Books sold under the traditional wholesale model. However, just using an OLS estimation approach would very likely lead to a biased estimator since it was probably not a random process whether an e-Book is sold under the agency or the wholesale model so that there might be a selection bias. Thus, we decided to apply a propensity matching design in which we identify appropriate treated and control e-Books through a matching procedure

[1] See United States v. Apple Inc., 12 Civ. 2826 (DLC).

(Rosenbaum & Rubin, 1983; Rubin, 1977). The results of the paper suggest that e-Books sold under the agency model are significantly cheaper than e-Book sold under the wholesale model. This outcome is also in line with many theoretical papers studying the price effect of agency arrangements, e.g., due to the elimination of double marginalization when using this type of contracts (Lu, 2017).

The third paper *Pass-through of temporary fuel tax reductions: Evidence from Europe* studies the pass-through of temporary tax reductions in the European gasoline market. In 2022, several European countries have implemented temporary fuel tax reductions to tackle very high inflation rates caused by ongoing supply chain problems following the COVID-19 pandemic as well as the war of aggression in the Ukraine. This paper studies the pass-through rate and the effect on the retail margins of those fuel tax reductions in the three largest countries of continental Europe (measured by GDP), namely France, Italy and Germany. Since the individual tax reductions have taken place at different points of time, a staggered Difference-in-Differences (DiD) design is employed to causally estimate pass-through rates and changes in retail margins (Callaway & Sant'Anna, 2021). Thereby, the paper relies on a unique panel data set containing consumer prices for gasoline and diesel on service station chain level. The results of this paper imply a full-shifting of the temporary fuel tax reductions meaning the estimated average pass-through rates are close to 100%. Nevertheless, in an event study design we find that the pass-through rates over time are heterogeneous between the countries and the type of fuel. In line with the estimated pass-through rates, no significant effect different from zero on the average retail margins in the three countries have been found. Important implications of these results for the effective design of unconventional fiscal policy as well as for competition policy in the fuel market are also discussed.

The following two papers are located in the field of health economics. Both papers empirically analyze COVID-19 vaccination campaigns and how different policies can enhance the efficiency of those campaigns. The World Health Organization (WHO) officially declared the previous COVID-19 epidemic a global pandemic on March 11, 2020 and within the last four years there have been almost 7 million fatalities in connection with this disease worldwide.[2] Vaccination campaigns have been considered a central pillar in the effort to stop this pandemic (Mallapaty, 2021). In the fourth paper *Efficiency in COVID-19 Vaccination Campaigns - A Comparison across Germany's Federal States*, the efficiency of the initial vaccination campaigns in the 16 German federal states are evaluated. Using data on vaccine deliveries and vaccinations given in Germany, we find considerable regional differences in the efficiency of the vaccination roll-out. In the second part of this paper, it is analyzed how the integration of doctors' offices into the campaign affected the speed of that campaign. The results of the paper imply that this integration significantly increased the efficiency of the vaccination roll-out. These findings are important for policymakers to identify best practices that can be adopted to speed up vaccination roll-outs in a country.

The fifth paper *Health Communication and COVID-19 Vaccine Hesitancy: A Synthetic Control Approach* analyzes the impact of health commu-

[2] See https://data.who.int/dashboards/covid19/deaths?n=c (data as of January 26, 2024).

nication on the willingness of individuals to get vaccinated against COVID-19. In order to answer this research question, a quasi-experimental design based on a speech by the President of France, Emmanuel Macron, is used. In this televised address, the French President tried to exert his influence by emphasizing the importance of getting vaccinated and also announced potential future restrictions for unvaccinated citizens.[3] The paper studies the impact of this speech on the daily vaccination rate in France by using the synthetic control method (Abadie & Gardeazabal, 2003; Abadie et al., 2010). Thereby, 13 other European countries are considered as a reference group in the donor pool. The analysis reveals that the speech by Macron has significantly increased the willingness of French citizens to get vaccinated against COVID-19. Hence, this finding suggests that applying leadership communication can be an effective weapon to change the beliefs of unvaccinated citizens and can possibly avoid the necessity of a general mandatory vaccination.

[3] See https://www.economist.com/europe/2021/09/18/how-france-tackled-vaccine-hesitancy.

2 The Substitutability Between Brick-and-Mortar Stores and e-Commerce – The Case of Books

Authors: Georg Götz, Daniel Herold, Phil-Adrian Klotz and Jan Thomas Schäfer

Presented at the following conferences:

- 46. Annual Conference of the European Association for Research in Industrial Economics (EARIE), Barcelona 2019

The Substitutability Between Brick-and-Mortar Stores and e-Commerce – The Case of Books

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

January 24, 2024

Abstract

We analyze competition between the online and the offline retail channel by using data on the German book market, which is characterized by fixed book prices. The analysis sheds light on the extent to which consumers perceive e-Commerce and traditional brick-and-mortar stores as substitutes. We find that, on average, when a bookstore closes, sales of print books decrease by around 744 units per month. This explains about on third of the total loss in sales of print books in our sample. These findings indicate imperfect substitutability between the online and the offline retail channel. Substitutability between the channels remains imperfect when we incorporate information on e-book sales. The magnitude of the effect is genre-dependent. For instance, sales of fiction titles decrease more strongly than sales of school books.

COI statement:

The data used in our study was provided by media control GmbH, GfK GmbH and Acxiom Deutschland GmbH. We received funding from the German Publishers and Booksellers Association ("Börsenverein des Deutschen Buchhandels e. V.") to buy the data.

*Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen, Germany.
We thank the German Booksellers' Association, Joel Waldfogel, Samuel de Haas, David Finck and Maximilian Maurice Gail for continuous support and fruitful discussions.
†Corresponding author. E-mail: daniel.herold@wirtschaft.uni-giessen.de.

Moreover, the German Publishers and Booksellers Association funded a research project at the Chair of Georg Götz from 2018 to 2020. Our study is a product of this project. The project itself was scientific in nature (i. e., no commercial research project). To conduct this project, the positions of the co-authors Jan Thomas Schäfer and Daniel Herold were funded by the German Publishers and Booksellers Association during the aforementioned period.

Keywords: Product differentiation, Book market, Retailing, e-Commerce
JEL Codes: L13, L81, D12, L42

1 Introduction

With the rise of e-Commerce, conventional brick-and-mortar retail sectors have experienced a substantial increase in competitive pressure across various industries (see, e.g., Burt and Sparks (2003) and Srinivasan et al. (2002)). It remains an open question to what extent and in which industries consumers perceive e-Commerce and traditional, physical stores as substitutes (see, e.g., Brynjolfsson et al. (2009), Sinai and Waldfogel (2004), and Wang and Goldfarb (2017)). In this article, we investigate the substitutability between e-Commerce and physical retailers in the German book market.

Physical bookstores and e-Commerce apparently have different "service" features. E-Commerce typically offers convenient product search options among potentially huge inventories (see, e.g., Waldfogel (2017)). However, ordering online entails waiting costs and e-Commerce does not offer the possibility of physical inspection (Guo & Lai, 2017; Loginova, 2009). Books are considered experience goods, so that pre-purchase uncertainty potentially affects demand (Hilger et al., 2011; Reinstein & Snyder, 2005). Against this background, consumers have the possibility to receive advice at physical bookstores or via online reviews or ratings (Reimers & Waldfogel, 2021).

Gilbert (2015) states that while reducing the importance of physical stores, e-Commerce might expand overall demand for books by attracting new customers (see also Cowen (2008)). However, given the differences in services offered by the two channels, it is unclear to what extent consumers view the offline and online channels as substitutes. If they are considered relatively far substitutes by sufficiently many consumers, it could be that a decrease in the number of physical bookstores (potentially triggered by advancing digitization (Gilbert, 2015)) decreases sales when (potentially captive) consumers of the offline channel reduce their demand for books. Analyzing the substitution patterns between the two retail channels can thus improve the understanding of how digitization affects the retail landscape. The German book market is particularly well-suited for such an analysis because it is characterized by fixed book prices such that the price-dimension plays no role in the comparison between the two channels.

To analyze to what extent consumers perceive the online and offline channels as substitutes, we investigate how closures of physical bookstores affect book sales. A novel panel data set is used which consists of monthly sales data of physical retailers and e-Commerce in Germany covering the period $2011-2017$. The data set also contains information on the number of physical retailers. The relationship between the number of physical bookstores and book sales may be bi-directional, i.e., demand may increase in the number of outlets or the number of outlets may increase in demand. We employ an instrumental variable (IV) approach using a proxy for population as an instrument to obtain consistent estimates.

Our findings indicate that, on average, market exit of one physical bookstore leads to a decrease in book sales of 744 units per month and federal states. This loss occurs in every subsequent month after the store's exit. Given that 1,382 bookstores were closed between 2011-2017 in our data set, the average loss in sales accumulates to around 43 million units. Given that the total loss in sales is around 137,000 million units, our results indicate that the closure of bookstores explains approximately on third of the decrease in total book sales.

The drop in sales appears to be genre-specific. For example, sales of fiction titles decrease by around 208 units on average per month and federal state when a book store closes, whereas we find no significant effect for schoolbooks. Remarkably, for fiction titles and children books we find that a decrease exit of physical bookstores implies also a weak decrease in sales in e-Commerce. This finding indicates complementary between the two channels in the sense that e-Commerce benefits from the presence of physical bookstores. This result can be considered evidence for a free-riding effect á la Telser (1960), as the existence of one retail channel affects the sales of another channel. Overall, our findings provide evidence that e-Commerce is no perfect substitute to physical bookstores.

We contribute to the literature on the impact of digitization and e-Commerce on "traditional" physical retailers.[1] For instance, Goldmanis et al. (2010) investigate how the presence of e-Commerce affects the competitive landscape in three markets in the US: travel agencies, bookstores and new car dealers. With respect to bookstores, they document a shift of market shares from small bookstores to larger chains in markets that are exposed to a larger degree of online competition, leading to a decline in the number of bookstores and employment. Goolsbee (2001) analyzes competition between online and offline retailers in the market for computer equipment. He finds that cross-price elasticities exceed unity and concludes that competition between the two channels is fierce. In a similar vein, Prince (2007) finds a strong jump in cross-price elasticities between online and offline re-

[1]See Section 2 for a more detailed discussion of the related literature.

tailers' computer products from 1996 to 1997. He traces this pattern back to demand- and supply-side effects such as online shopping becoming more prominent as well as an expansion of services of online retailers. In the context of the impact of digitization on the book market, our paper is closely related to Brynjolfsson and Smith (2000), Clay et al. (2002) and Chevalier and Goolsbee (2003). These studies analyze price dispersion and demand elasticities and focus on markets where retailers compete in prices. In the absence of price differences between the online and the offline channel, we find that the decrease in sales triggered by the closure of a bookstore are not fully compensated by increased sales of other bookstores or in e-Commerce. As an exit is equivalent to an infinite increase in prices in the respective bookshop, we conclude that the resulting decrease in sales at the exiting store is not compensated one-for-one either by sales in other physical bookstores or in the online channel. The obvious reasons are transport and inconvenience costs, respectively. Our findings therefore indicate that physical bookstores generate additional sales. This also holds for the online channel. Apparently both channels contribute to welfare.

We also contribute to the literature and policy discussion on vertical restraints, particularly with respect to the evaluation of potential efficiency effects arising from resale price maintenance (RPM). This evaluation has become an important topic of competition policy at least since the *Leegin*-Case, which led to RPM being shifted from per se illegality to a rule of reason based approach in the US in 2007. In the EU, RPM is still practically considered per se illegal (see Akman and Sokol (2017) for a more detailed overview). However, in various European countries such as France, Austria and Germany (and other countries such as Japan)[2], fixed book price systems are in place, which occur in the form of RPM. As the goal of fixed book prices is usually to protect books as a cultural or merit good, they constitute an exception when it comes to the application of competition law.[3] Whether such an exemption from some fundamental pillars of competition law is justified, crucially depends on whether fixed book prices can be considered an appropriate tool to achieve the policy goal to support the book market, i.e., to increase the demand and supply of books. Against the background of our findings, one way through which fixed book prices could affect book sales would be securing margins and thereby potentially promoting market entry or preventing exit of physical retailers (Bouckaert, 2000; Elzinga & Mills, 2008; Marvel & McCafferty, 1985).

The article is structured as follows. In Section 2, we review the related literature. This Section also contains a brief conceptualization of our empirical analysis based on economic theory. Our data set is described in Section 3. Section 4 presents our empirical analysis. Section 5 provides an overview of the robustness checks that are presented in the Appendix. Section 6 concludes.

[2]See, for instance, Global Fixed Book Price Report of the International Publishers Association, May 23, 2014, https://bit.ly/2Wg1tpz. According to Poort and van Eijk (2017), in 15 OECD-countries (ten EU-members) there is a fixed book price system in place.

[3]The German law on fixed book prices ("Buchpreisbindungsgesetz") lays down the specifities and goals of the regulation, see https://bit.ly/3qU6K4b.

2 Related Literature & Theory

2.1 Literature Review

The present paper is closely related to the literature on competition between the online and the offline channel. In addition to the articles discussed in the introduction, there are numerous articles in this field such as Goolsbee (2000), Brown and Goolsbee (2002), Jin and Kato (2007), Ofek et al. (2011), Gauri et al. (2021) or Couture et al. (2021). In the following, we focus on the most closely related papers that were not already discussed in the introduction. In addition to discussing the related literature, we emphasize some key service elements in which the online and offline channel differ. This step will be necessary to substantiate our theoretical and empirical analyses that follow in subsequent sections.

The impact of digitization has been studied from different angles, including the topics crowd ratings (Reimers & Waldfogel, 2021), the role of niche-products (Brynjolfsson et al., 2003, 2006; Reimers & Waldfogel, 2017) and innovation, diffusion and copyright protection (Reimers, 2016; Waldfogel, 2017; Waldfogel & Reimers, 2015). Most studies on books have so far focused on markets where retailers compete in prices.[4] For example, Brynjolfsson and Smith (2000) compare prices of books and CDs between online and offline retailers. They find that 33% of the considered books are significantly cheaper in the online channel and that there is substantial price dispersion in that channel. Clay et al. (2002) partly confirm these observations. In their sample of 107 books, they find price dispersion online, however, prices

[4]One exemption would be Beck (2007), who analyzes the effects of word-of-mouth.

in the online and the offline channel were similar. Chevalier and Goolsbee (2003) use sales ranks for books to estimate price elasticities and calculate price indices for the online retailers Barnes and Noble and Amazon.

Some markets are characterized by a large degree of pre-purchase uncertainty regarding features, effectiveness or quality. This uncertainty can give rise to various inefficiencies (Bagwell & Riordan, 1991; Balasubramanian, 1998; Bester, 1998; Legros & Stahl, 2019; Riordan, 1986). Books appear to be no exception, especially because they are considered experience goods (Nelson, 1974), which means that the uncertainty prior to purchase is pronounced. We consider the reduction of this uncertainty as one key feature of service provision of the two channels, in which they differ substantially. This service provision can be considered an important dimension of competition between the channels (Gilbert, 2015).

Even though we cannot isolate which factor is pivotal for consumers, our paper is related to the aforementioned strand of literature. A key feature here can be the role of expert opinion to reduce the consumers' uncertainty (Clement et al., 2007; Lizzeri, 1999; Reimers & Waldfogel, 2021). Expert opinion might affect demand in various ways. For instance, consumers may learn that a certain type of good exists (Hilger et al., 2011). Moreover, the opinion can act as a signal of quality (Reinstein & Snyder, 2005). Thus, 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).[5] A reputable book dealer in a nearby

[5]It is arguably the case that the shelf space in physical bookstores is limited, whereas in e-Commerce this space is "limiteless" (Rabinovich et al., 2011). However, with respect to the German book industry, it has to be noted that every book that is on the list of

physical bookstore may also offer expert opinion. Closely related is the so-called word-of-mouth effect as analyzed by Beck (2007). In e-Commerce, such as on *Amazon.com*, online reviews written by other readers are presented and books are recommended to the customer based on the same customer's or other readers' past purchases, which can also be considered a form of service provision or expert opinion (see, e.g., P.-Y. Chen et al. (2004)).

Hilger et al. (2011) find empirical support for the hypothesis that expert opinion may have a positive effect on the demand. They show that expert opinion helps customers to find a desired wine, i.e., demand for wine whose quality is considered low (high) by experts are purchased less (more) often. Hilger et al. (2011) find that there is no one-for-one substitution pattern, so that overall demand increases when expert opinion is available. This may be due to existing consumers purchasing not only wine whose quality they perceive as high but also additional wine or due to new consumers who would not have bought wine in the absence of expert opinion. This finding is supported by Mérel et al. (2021), who attribute a 7% loss in welfare to asymmetric information in the French wine market, which corroborates the importance of certification schemes. In another related paper, Reinstein and Snyder (2005) empirically analyze the effect of expert movie reviews on box office revenue. For narrowly-released movies and dramas they find a significant correlation between expert reviews and box office sales whereas for widely-released movies and other genres such as action movies they do

available books (the so-called "Verzeichnis lieferbarer Bücher") can be ordered in physical bookstores as well so that essentially the same books are available in both channels because they are served by the same wholesalers. If ordered offline the book will be available the next work day. See https://bit.ly/3Pdvu2j.

not find a significant correlation. In this context, Reimers and Waldfogel (2021) demonstrate that crowd-based ratings such as the commonly used star rating schemes have increased consumer surplus over ten times more strongly than the traditional system of professional critics in the book market. This indicates that online information provided by other consumers is valuable from a welfare perspective.

A distinctive feature of e-Commerce is consumer search. To name only a few examples, it has proven effective in various instances such as the job market (Gürtzgen et al., 2021), sport card trading (Jin & Kato, 2007), the insurance market (Brown & Goolsbee, 2002) as well as the book market (e.g., Goldmanis et al. (2010)). The role of search for competition policy has, for instance, been discussed in the context of price parity clauses (Mantovani et al., 2021). Tang et al. (2010) demonstrate that the rise of the use of shop bots as a tool of consumer search has decreased prices and price dispersion for books. Thus, the ability to conveniently browse through potentially millions of books online (Waldfogel, 2018) and find a specific title is considered a distinctive service element of e-Commerce.

Another dimension in which the online and the offline channel might differ is ad hoc sales. The number of impulse purchases in the offline channel may be higher (see, e.g., Burt and Sparks (2003)). In that respect, ordering online entails waiting costs while purchasing offline allows the consumer to immediately receive the good (Bouckaert, 2000; Loginova, 2009).

Physical retailers also have the opportunity to provide showrooms where consumers can physically investigate a product (Jin & Kato, 2007; Loginova, 2009). This might positively influence demand in some instances (Bell et al., 2017). Even though it might not seem important how books look like, note that you can hire "book curators" who are specialized in how to make your bookshelf look better.[6] Physical inspection, however, requires that a consumer travels to the next bookstore.

In terms of analyzing consumers' substitution patterns between sales channels in the book market, this article is also loosely related to the literature on demand estimation, which has become a central part of empirical economics since the pioneering works of Berry (1994) and Berry et al. (1995). In this article, we focus on substitution patterns between retail channels rather than between differentiated products (i.e., books). Given that prices are the same in both channels due to Germany's fixed book price regime, we do not use a structural model such as (random coefficient) nested logit.

Finally, there are theoretical papers analyzing competition between the offline and the online channel. Related papers include Balasubramanian (1998), Bouckaert (2000), Chu et al. (2012), Guo and Lai (2017), and Loginova (2009) or Legros and Stahl (2019). Guo and Lai (2017) is particularly closely related to our approach. In their model, physical retailers and e-Commerce compete for consumers located in rural and urban areas (i.e., for non-uniformly distributed consumers along the Hotelling-line). They show that in the short-run, when exit is not possible, competition decreases prices. However, prices increase in the medium-run, when retailers can exit the market. In the long-run, the surviving physical retailers relocate to urban areas. The model of Chu et al. (2012) emphasizes competition between online and

[6]See, for instance, an article from the publisher Penguin's homepage from May 06, 2020, https://bit.ly/3t3Gw1M (last accessed October 12, 2022).

physical bookstores. They assume that e-Commerce offers a larger variety, but ordering online entails waiting costs. In a similar vein, Legros and Stahl (2019), building on previous models in Stahl (1982) and Schulz and Stahl (1996), analyze the role competition from e-Commerce (external competition) plays with respect to competition between local physical retailers (internal competition) in a recent working paper.[7] Consistent with Guo and Lai (2017), the authors show that more intense external competition reduces variety in the internal market.

2.2 Theoretical Background

The online and offline retail channels have different features, as explained in Section 2. Consumers vary regarding their preferences towards these features. To grasp this concept formally, one could think of consumers' choice in a setting of (horizontal) product differentiation. Indeed, these kinds of problems are formalized in the literature using Hotelling- oder Salop-models where consumers' preferences regarding the two sales channels are reflected in different transport costs.[8] For instance, in Balasubramanian (1998), Bouckaert (2000), Chu et al. (2012) and Guo and Lai (2017) it is assumed that each consumer also has the opportunity to buy online. In these models, inconvenience costs are assumed to be the same across all consumers, which basically means that e-Commerce is located in the center of a circle in a Salop-framework.

[7] Legros and Stahl (2019) constitutes a revised version of an older working paper from 2002.

[8] An alternative approach was chosen by Loginova (2009). She formalizes waiting costs in e-Commerce by deflating consumers' maximum willingness to pay.

Conceptually, we analyze how exit of a store affects demand, or, in other words, the substitutability between physical stores and e-Commerce. As mentioned above, this has been and could be formalized in a Salop-circle or Hotelling-line.[9] Such a formalization, however, would require a model where not all consumers always purchase the product. Technically, the model would have to allow for partial market coverage (instead of always having full market coverage). In order to formalize this, one would need to depart from the assumption of symmetric inconvenience costs from buying online by assuming that these costs differ between consumers.

To understand why this is necessary, it proofs convenient to adopt two illustrative terms established, e.g., by Somogyi (2020): the market splitting condition defines consumers being indifferent between two retailers (or, more generally, variants) and the participation condition defines which consumers are willing to buy the good. Assuming symmetric inconvenience costs across all consumers is appropriate to analyze competition between the two channels, which will be mainly driven by market splitting conditions. (The participation condition will have to be satisfied as well in equilibrium, since otherwise demand would be zero.) However, our goal is to examine the influence of exit of physical bookstores on sales. In that sense, the participation condition will be of major importance. When all consumers share the same inconvenience costs from buying online and prices are fixed (as it is the case in the German book market), the participation condition in e-Commerce will be the same for every consumer. That is, there is either always full market

[9] An earlier version of this paper contained such a model. It is available from the authors upon request.

coverage (as every consumer would buy online in the absence of physical bookstores) or no consumer will buy online.

In order to investigate how market exit of a physical bookstores can affect sales, one thus needs to depart from the assumption that every consumer has the same inconvenience costs from buying online.

It is possible in such a setting that exit of physical bookstores *ceteris paribus* leads to a decrease in market coverage and, thus, demand for books. We interpret this pattern as *imperfect substitution* between the online and the offline channel. That is, some consumers who patronized a physical bookstore and who have high inconvenience costs from buying online may stop buying altogether because neither e-Commerce nor the remaining bookstores are perceived as "close enough" substitutes. This notion may serve as a theoretical interpretation of the following empirical analyses.

3 Data and Descriptive Statistics

For our empirical analysis, we use monthly data on the number of book retailers as well as sales volumes and revenue data on a federal state level. The data comprise sales of books in brick-and-mortar stores as well as e-Commerce sales, e.g. books shipped by Amazon. We distinguish between nine different product groups to control for genre-specific effects. The database covers the period from January 2011 to December 2017. We utilize a combination of data from multiple sources. In what follows, we present some descriptive statistics and we describe the data sources in more detail.

3.1 Brick-and-mortar bookstores

The German Publishers and Booksellers Association maintains several databases on the number of brick-and-mortar stores that we've combined in order to obtain an accurate picture of the market and its' development over time.[10] Figure 1 shows the development of the number of physical bookstores in Germany over time, based on data from the German Publishers and Booksellers Association. The data includes information on outlets of chain stores as well es independent bookstores. Between 2011 and 2017, the number of physical bookstores decreased from over 6,300 to less than 4,900.

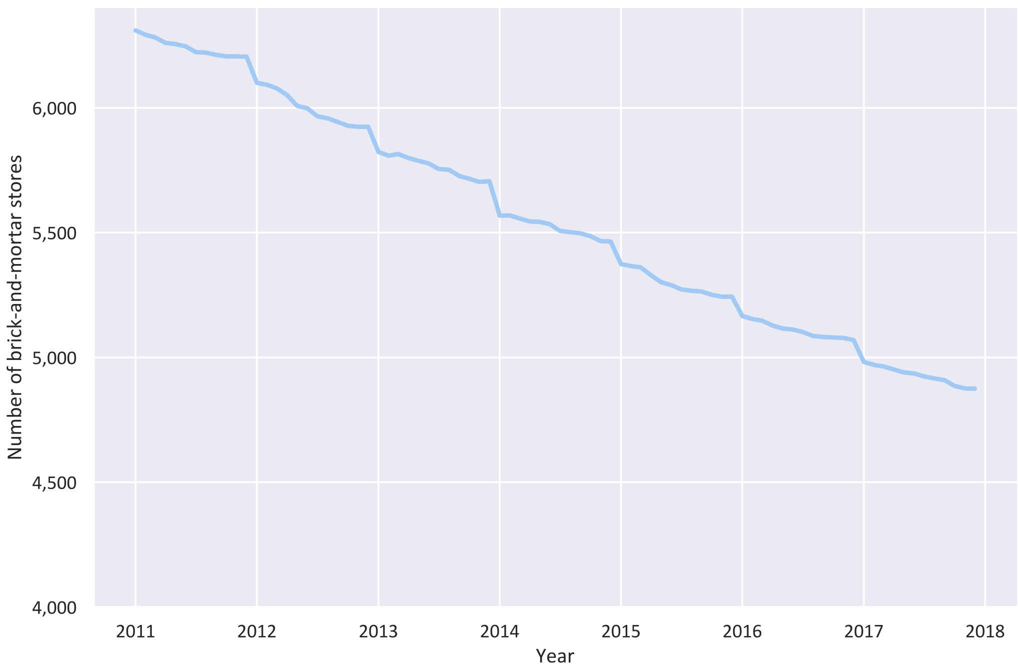

Figure 1: Development of physical bookstores over time. Source: German Publishers and Booksellers Association.

The numbers presented in Figure 1 are derived from the membership database of the German Publishers and Booksellers Association. Roughly

[10] Börsenverein des Deutschen Buchhandels, see https://www.boersenverein.de.

$85 - 90\%$ of all book retailers in Germany are members of the German Publishers and Booksellers Association. The database contains information on the name, the address, the founding year of a bookstore, the date of the beginning as well as the end of a membership in the Association. Since 1963 a unique identifier ("Verkehrsnummer") is assigned to each business entity that is member of the German Publishers and Booksellers Association. This identifier is used for communication between business partners, i.e. for communication between bookstores, publishers, and wholesalers. Using this identifier we are able to match bookstores to other databases, like the so-called Address Book for the German-Language Book Trade.[11] This address book was established in 1839, and is now managed by a subsidiary of the German Publishers and Booksellers Association.[12] The database contains over 30,000 addresses of publishers, bookstores, music stores, and publishing representatives in the German-speaking region. The address book is updated regularly and changes (entry, exit, change of location or legal name) are published online. Members of the German Publishers and Booksellers Association can be listed in the address book for free, non-members have to pay a small annual fee.

We apply several steps in order to transform the merged address book/ membership database into a monthly panel that contains the number of active brick-and-mortar stores for each Federal State. First, we drop all publishers, head offices and warehouses as well as online bookstores from the database.

[11] Adressbuch für den deutschsprachigen Buchhandel, see https://adb-online.de; last accessed December 19, 2023.

[12] MVB Marketing- und Verlagsservice des Buchhandels GmbH, see https://mvb-online.de; last accessed December 20, 2023.

Second, we obtain information on the Federal State in which each brick-and-mortar store is located, using the Google Maps API. Third, we group stores by location. In doing so, we avoid situations where a bookstore would receive a new identifier due to, e.g., a change in ownership-structures, despite operating at the same location. Fourth, we determine the opening and, if possible, closing date of each group and transform the data base to a panel structure by counting all active brick-and-mortar stores for all federal states and months.

We checked the consistency of our data set using data on the order volume of bookstores provided by three of the largest German wholesalers.[13] Using the unique identifier we were able to match the order data with the address book/membership database. There are only few bookstores that are customers of the wholesalers but are not member of the German Publishers and Booksellers Association, and accordingly do not have a unique identifier. These relatively small bookstores might be missing in the address book/ membership database. It should be noted that it seems likely that these booksellers do not report their sales data to market research organisations at all, due to their small size. Thus, we do not expect that our analyses are distorted by missing information for some bookstores.

Thus, we believe that our data set provides a reliable picture of the developments in the German book market. However, there is some uncertainty regarding the exact timing of entry or exit from the market. This is attributed to the design of the reporting mechanism of the address book/ membership

[13]Libri GmbH (https://www.libri.de)G. Umbreit GmbH & Co. KG (https://www.umbreit.de) and Koch, Neff & Volckmar GmbH (KNV, http://www.knv.de).

database. While the data on entry appears to be reliable, as new bookstores are likely to obtain a unique identifier to expedite the order and billing process as quickly as possible, the date of market exit may not necessarily be exact. Exiting bookstores may not communicate the end of their business operations as promptly as possible or not at all. In those cases the members are removed from the database if a default in payment of membership fees is determined.

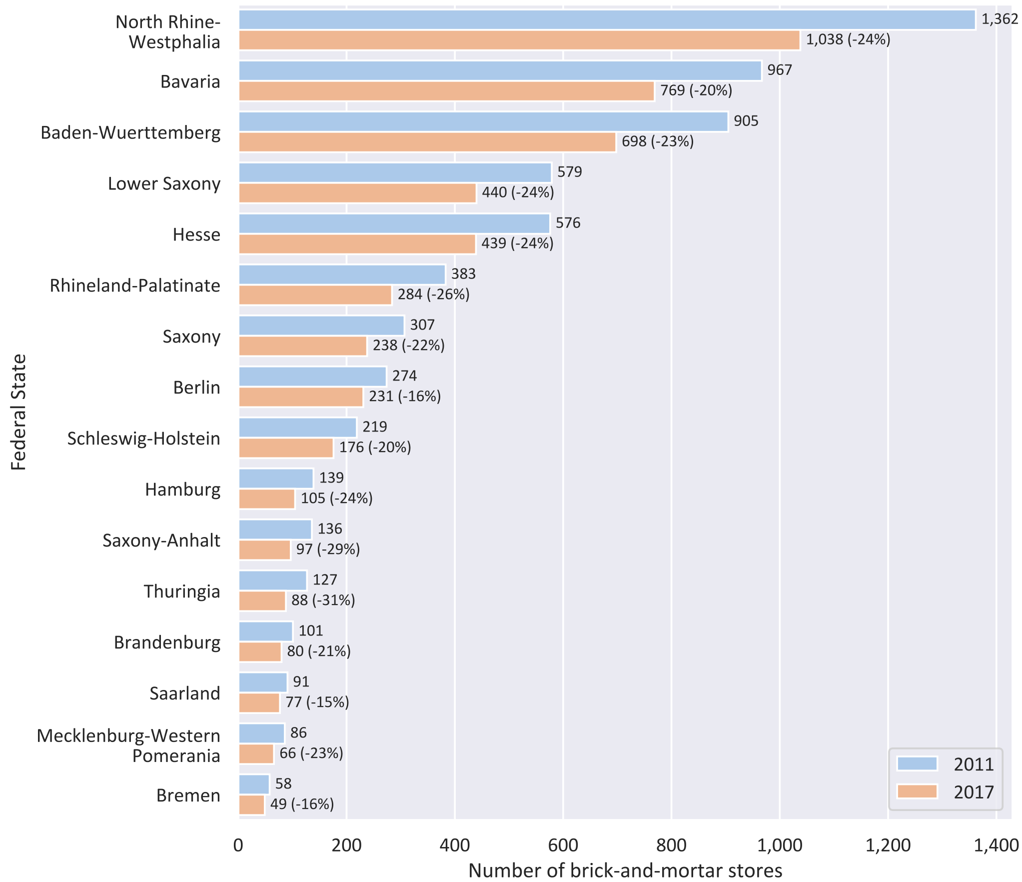

Figure 2: Development of physical bookstores on the federal state level. Source: German Publishers and Booksellers Association.

Figure 2 provides summary statistics of bookstores on the federal state level. On average, the number of brick-and-mortar bookstores decreased by 20 percent over the observation period. In absolute numbers, around 1,435 bookstores exited the market. Figure 3 shows the number of brick-and-mortar stores per 100,000 residents in beginning of 2011 and end of 2017.[14]

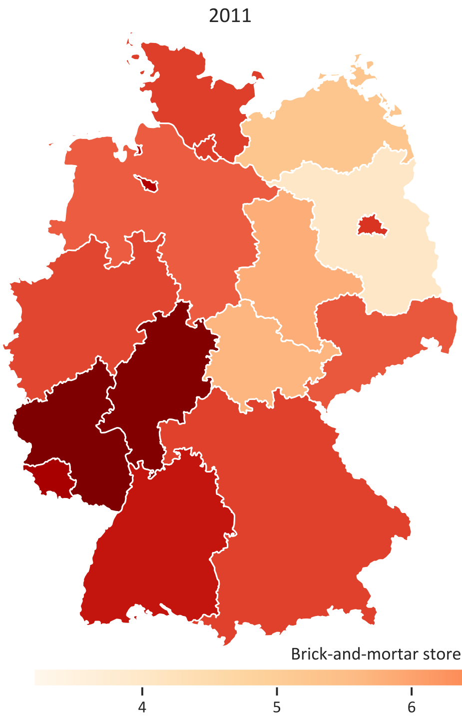

Figure 3: Number of physical bookstores per 100,000 residents in beginning of 2011 and end of 2017. Source: German Publishers and Booksellers Association.

3.2 Sales data

In order to study the development of sales over the observation period, we use data from two sources. First, we have access to scanner data provided by media control GmbH.[15] Second, we use consumer panel data provided by

[14]Note that we use labor force as a proxy for population. This is because population data is only published on a yearly frequency. See Section 4 for further discussion.

[15]see https://www.media-control.de/buch1.html; last accessed December 20, 2023.

GfK GmbH.[16] Both databases differ with respect to coverage, frequency and accuracy. The scanner data is based on quasi real time data and comprises information on sales of print books in independent bookstores, chain stores as well as online retailers (see Appendix A.1 for more characteristics of the data set). The data provider claims to cover more than $80-90\%$ of the German print book market. However, the scanner data do not contain information on the sale of e-Books. To the best of our knowledge, the most reliable information on e-Book sales in Germany available to researchers are consumer panel data, which, for the observation period 2014-2017, is provided by GfK. The sales data projections of GfK are based on roughly 20,000 consumers being surveyed on a regular basis and contain information on the purchases of print books, e-Books and audio books. In contrast to the panel data, the survey data is only available on a quarterly level for the period from Q1.2014 to Q4.2017 (see Appendix A.2 for more characteristics of the data set). For the years 2011 to 2013 survey data are only available as yearly aggregates and for whole Germany, rather than on the federal state level and are included for illustrative purposes only.[17]

Figure 4 shows sales figures for physical books from the scanner data set. It becomes obvious that the sales of print books in Germany exhibit a decreasing trend. In 2011, 297.2 million print books were sold in brick-and-mortar and online stores with a corresponding revenue of 3.66 billion Euros. This implies an average book price of 12.31 Euro. In 2017, 264 million books

[16] see https://www.gfk.com/de; last accessed December 20, 2023.

[17] For e-Book sales see https://de.statista.com/statistik/daten/studie/232191/umfrage/absatz-von-e-books-in-deutschland; last accessed December 20, 2023. For total sales (including e-Books and audio books) see https://de.statista.com/statistik/daten/studie/416380/umfrage/absatz-von-buechern-in-deutschland; last accessed December 20, 2023.

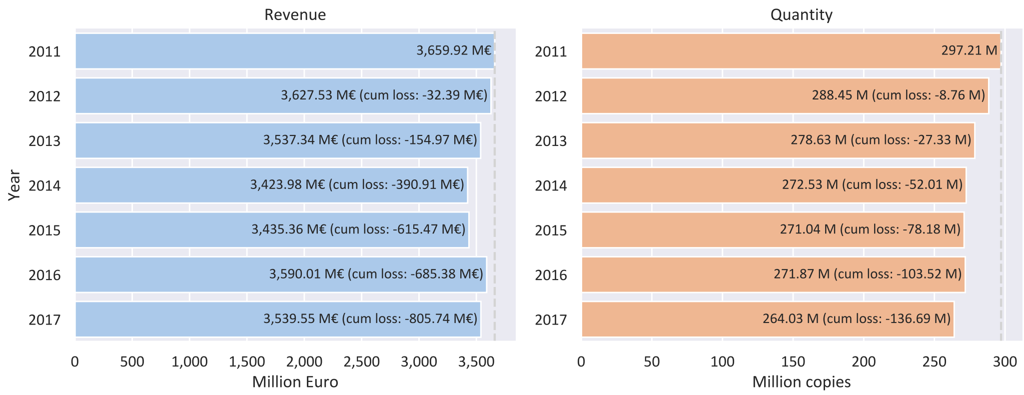

Figure 4: Annual revenues (in million Euros) and sales (in million books sold) of print books in brick-and-mortar stores and e-Commerce. Cumulative losses since 2011 in parentheses. Thuringia and Saarland are excluded due to inconsistencies in the data. Source: media control GmbH.

were sold with a revenue of 3.54 billion Euros (i.e., 13.41 Euro per book). The decline in sales was thus at least partly offset by a price increase so that revenues of brick-and-mortar and online stores remained fairly constant.

As can be seen in Figure 5, the number of print books sold per resident decreased between 2011 and 2017.

At the same time, e-Book sales increased from 4.30 million in 2011 to 29.15 million in 2017. As can be seen from Figure 6, the growth rate of e-book sales stalled after a sharp increase and corresponding revenues even declined between 2016 and 2017. It is important to note that e-Book sales data cannot be directly compared to the data on print book sales that have been presented in Figure 4. From the comparison of GfK's print book sales data (see Appendix A.2) with the more accurate scanner data, we know that the consumer panel projections might be overestimating ”true” sales by up to 25 percent. However, even without correcting for this potential bias, it becomes clear that e-Book sales did not completely compensate the decline in sales of print books in all years. For example, print book sales

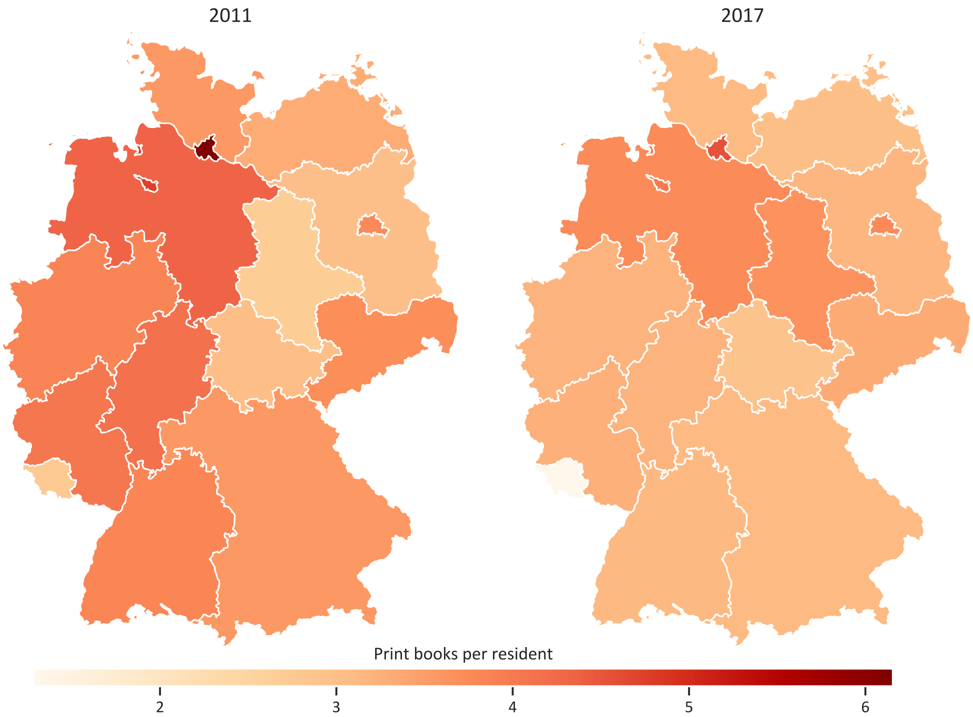

Figure 5: Number of print books sold per resident and year. Source: media control GmbH.

dropped by 6.1 Million copies between 2013 and 2014, while e-Book sales only increased by 3.31 Million. Nevertheless, market shares of e-Books are increasing, however, they are on average not higher than 8 percent according to the consumer panel data.[18]

In what we presented so far, we have only looked at yearly sales aggregates. To illustrate developments during the years, Figure 7 shows monthly aggregates of print book sales in brick-and-mortar and online stores. A large share of the revenues in the print book market is generated in December. This particularly strong "Christmas effect" is widely acknowledged by scholars of the book market (see, e.g., Beck 2007). The sales and revenue patterns

[18]According to the German Publishers and Bookseller's Association the market shares of e-Books were around 5% in 2017 and 2018, see https://bit.ly/2TSJHas.

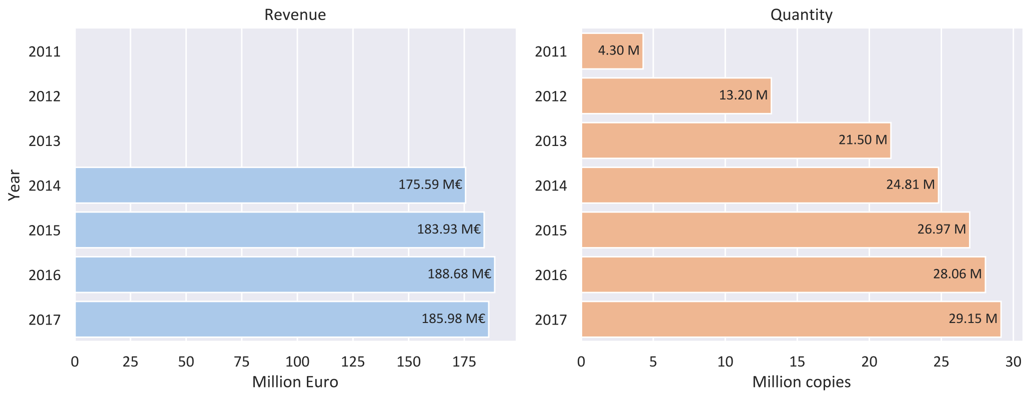

Figure 6: Annual revenues (in million Euros) and sales (in million books sold) of e-Books. Source: GfK GmbH.

are characterized by another smaller peak at the beginning of summer, when the new school year begins. Until 2014, a decline in sales in peak times can be observed. Between 2014 and 2016 revenues in peak times slightly increased again, while sales remained roughly constant.

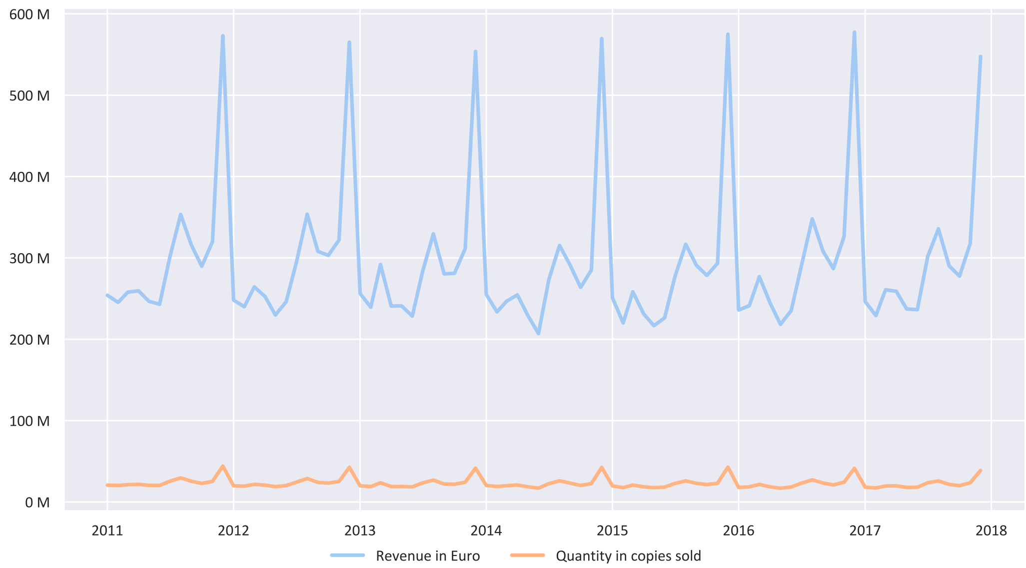

Figure 7: Aggregated sales data (print books) over time (month). Thuringia and Saarland are excluded due to inconsistencies in the data. Source: media control GmbH.

Figure 8 shows the development of monthly sales of print books in brick-and-mortar stores. In our sample, an average physical bookstore generates a revenue of 46,145 Euros per month by selling 3,760 books. That is, 45,120 print books are sold per year by an average physical bookstore.

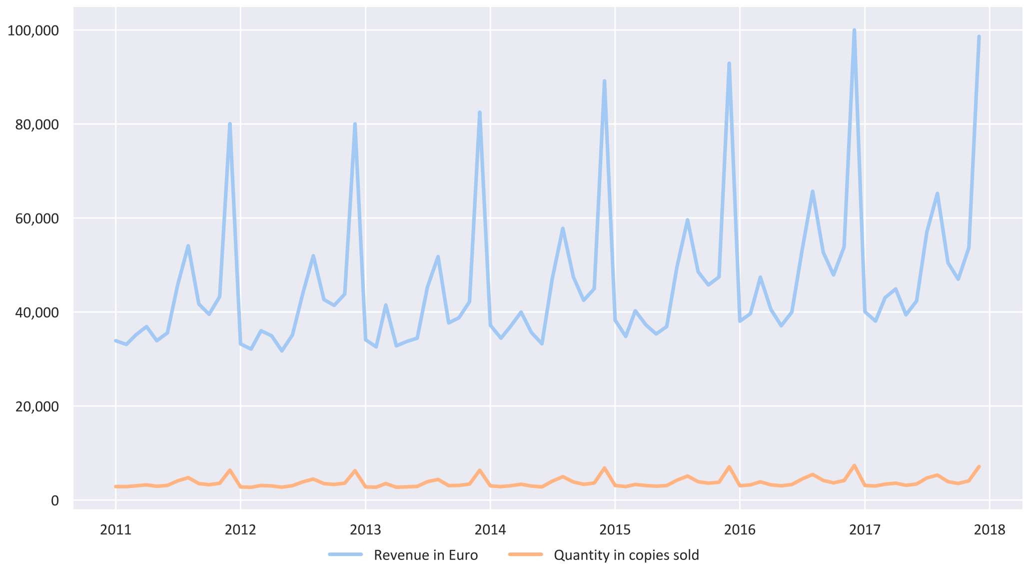

Figure 8: Print book sales per brick-and-mortar store (monthly mean). Thuringia and Saarland are excluded due to inconsistencies in the data.

Even though the average revenues of brick-and-mortar stores are increasing over time, market shares are declining. That is, print books are increasingly being ordered online. Between 2011 and 2017, the revenue-based market share of e-commerce increased from about 15 percent to almost 19 percent.[19] However, as we will demonstrate in the following sections, neither online stores like Amazon nor e-Books can fully compensate for all sales if a brick-and-mortar store closes.

[19]See German Publishers and Booksellers Association, Buch und Buchhandel in Zahlen, available online https://www.boersenverein.de/markt-daten/marktforschung/wirtschaftszahlen/; last accessed January 15, 2024. Please note that we can also determine market shares based on scanner data. However, due to confidentiality agreements signed by us, we are unable to disclose this information. The figures we have determined, however, match in terms of development and approximate magnitude with the values reported by the German Publishers and Booksellers Association.

3.3 Google Trends Indices

The empirical analysis will be conducted using two distinct datasets. First, we will employ the scanner dataset (Section 4.2), and, second, we will utilize the consumer panel dataset, which, while including e-Book sales, provides lower frequency and coverage (Section 4.3). In both cases, we will use panel regressions to analyze the effect of the closure of bookstores on sales. We will use two Google Trend Indices to account for e-Book sales and state-specific time trends. This is particularly important in the analysis of the scanner data because this data set does not contain information on e-Book sales. A more thorough explanation of our estimation approach will be presented below. In what follows, we briefly explain the aforementioned indices.

The first variable we use in our analyses is the Google Trends Index for Amazon Kindle (Topic E-Book-Reader). The second variable is the Google Trends Index for the topic 'Book'. Topics are generally considered to be more reliable for Google Trends data than using exact search terms. They are constructed based on exact phrases as well as misspellings and acronyms, and cover all languages.

Google Trends data is described to be drawn from a random, unbiased sample of Google searches. Note that exact numbers of the respective terms, topics or search queries the respective indices are based on are not available; only the indices are available. The index ranges from 1-100, where 100 is the maximum search interest over the whole observation period and all federal states.[20]

[20] See https://newsinitiative.withgoogle.com/en-gb/resources/trainings/basics-of-google-trends/; last accessed December 26, 2023

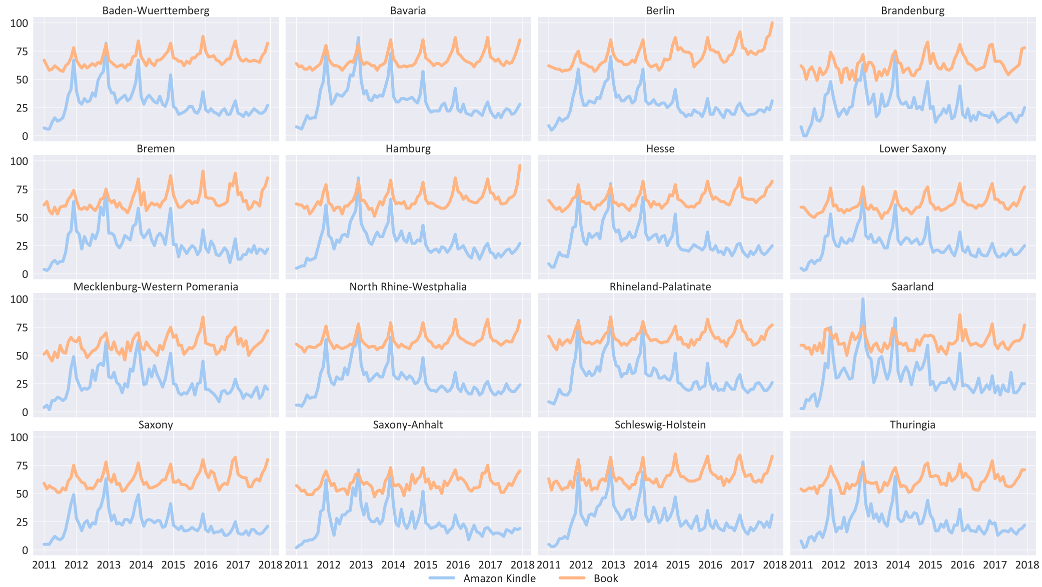

Figure 9: Google Trends Index for Amazon Kindle and the Topic Book.

The Index for the topic E-Book-Reader reflects search behavior for the term 'Amazon Kindle' and its acronyms that are directly related to e-book readers, such as searches for specific models, features, or comparisons with other e-readers. We consider the prevalence of online consumer search for e-Book readers to reflect their demand and, therefore, to act as a proxy for e-Book sales, since these readers are required to conveniently read e-Books. Thus, in our analyses, the trend index for E-Book-Reader is used to control for changes in e-Book demand over time. As can be seen by comparing the development of the index in Figure 9 with the evolution of e-Book sales shown in Figure 6, both are characterized by a sharp increase in the earlier years, while (relative) search interest decreases in the later years, and sales of e-Books remain rather constant. Similar to the sales pattern shown above, the search index also includes a "Christmas effect".

Figure 9 also depicts the development of the second index that is used in our analyses, the Google Trends Index for the topic 'Book'. This index includes search queries for the German word for book ("Buch"). It potentially reflects state-specific trends in reading behavior, hence we use it to capture variations in demand for books. The advantage of incorporating this index into our analyses is that it might capture state-specific time-variant fluctuations in online search and, therefore, demand for books. Such variations might, for instance, stem from differences in school holidays between federal states, which, in turn, might have an effect on demand for books in each federal state. The index therefore provides additional information and is used to complement our fixed-effects panel regression outlined in what follows.

4 Empirical Analysis

The goal of our empirical analysis is to analyze the influence of the number of physical bookstores on book sales. This relationship gives rise to concerns of reverse causality. On the one hand, the number of physical bookstores can affect sales. On the other hand, monthly sales reflect demand, which affects the number of retailers on the supply side of the market. To resolve this source of potential endogeneity, we will employ an IV approach using the labor force as an instrument in our estimations. This will be explained in more detail in Section 4.1 .

After explaining our IV strategy, we present the first step of our two-step analysis in Section 4.2. The estimations presented there focus on print book sales. In a second step, we incorporate e-Book sales into our analysis (see Section 4.3).

Before we discuss our IV approach in Section 4.1, we will first outline our identification strategy. If e-Commerce and brick-and-mortar stores were perfect substitutes, the drop in the number of physical stores should not have any impact on book sales because consumers will switch to the e-Commerce channel. A positive relationship between the number of physical stores and book sales thus indicates that e-Commerce and physical stores are imperfect substitutes. In the IV-approach, a coefficient that is not significantly different from zero indicates perfect substitutability as a loss of sales of a closed physical bookstore does not lead to a detectable decline in sales. With respect to e-Commerce sales, the sign of the coefficients have to be interpreted differently. A positive (negative) coefficient implies complementarity (substitutability) between the two channels because a lower number of physical bookstores leads to a decline (an increase) in sales. This will be discussed in more detail in the context of the respective output tables presented below.

The relationship between e-Books and print books is an important driver of the results. The closer the degree of substitutability between e-Books and print books the more likely it is that the decrease in sales of print books is compensated (or even over-compensated) by an increase in e-Book sales. Gilbert (2015) conjectures that e-Books are imperfect substitutes to hardcover books, which is empirically supported by Li (2015) and H. Chen et al. (2019). Li (2019) finds a high degree of substitutability between e-Books and paperback books. It is thus possible that there is substitution between physical books and e-Books when physical bookstores close.

4.1 IV–Approach

The goal of the following empirical analyses is to examine the impact of the number of physical bookstores on book sales. However, simply regressing book sales on the number of physical bookstores constitutes almost a textbook-example of endogeneity. First, the direction of effects is unclear – do fewer physical bookstores lead to fewer sales or is it the other way around? Second, there might be unobserved heterogeneity because sales and the number of physical bookstores may be driven by unobserved variables such as overall reading behavior.[21] To solve this problem, we employ an instrumental variable approach.

Technically, the instrument has to satisfy the following conditions. The dependent variable on the first stage is the number of bookstores. An instrument has to satisfy that the number of bookstores and the instrument must be correlated (relevance condition). The dependent variable on the second stage is booksales per capita. The instrument and book sales per capita must not be correlated (orthogonality condition). (See Wooldridge, 2015, p.463.)

Our instrument of choice is population. This rests upon the assumption that there is a correlation between population and the number of bookstores (relevance condition) and that population does not affect the dependent variable in the second stage, which is book sales per capita (orthogonality condition). These two conditions are explained in more detail in what follows.

First, consider the relevance condition. One would expect that, *ceteris*

[21]Note that time trends potentially affecting reading behavior on a national level such as the growing popularity of streaming services are controlled for in our estimations by time fixed effects.

paribus, in federal states with higher population there are more bookstores. In particular, we allow for a non-linear relationship between those two variables. This assumption appears valid against the background of our empirical results since the test statistics suggest that the potentially endogenous variable (number of bookstores) is not sufficiently identified to have an effect in the second stage when we only use the (linear) population size as an instrumental variable. The respective regression results of the first stage of the IV approach are shown in Appendix B.2. The first stage results imply that the population size has a significantly positive effect on the number of physical bookstores in a federal state, while the effect of the squared population variable is significantly negative. This result implies a non-linear (concave) relationship between the number of physical bookstores and the population size on federal state level.[22]

Next, consider the orthogonality condition. The dependent variable in the second stage of our estimation approach is book sales per capita. The orthogonality condition should be satisfied as there should not be a correlation between a federal state's population and the average number of books a person reads. The former is basically a demographic factor, whereas an individual's inclination towards reading should mainly be driven by personal preference or lifestyle. In other words, we expect that, *ceteris paribus*, in a federal state with 4 million inhabitants each individual reads the same number of books in every period as an individual in a federal state with 18 million

[22]To put this into perspective, on could think of an entry model. When market size increases, the number of firms in the market should also increase. A concave relationship between the number of physical bookstores and population size indicates that the number of additional entrants decreases in demand.

inhabitants.[23]

Formally, denote by $\#stores_{i,t}$ the absolute number of physical bookstores in federal state i and month t. Population is captured by $pop_{i,t}$, which is used to instrument for the potentially endogenous variable $\#stores_{i,t}$. As explained above, we allow for a non-linear relationship between the number of physical bookstores and the population in a federal state so that we also incorporate the squared population ($pop_{i,t}^2$) as an additional instrument.

The relevance condition requires that the instrument is correlated with $\#stores_{i,t}$.[24] The dependent variable in the second stage is book sales per capita, denoted by $sales\_pc_{i,t}$, for federal state i in month t. The orthogonality assumption requires that a change in population does not directly affect a change in book sales per capita, i.e. $pop_{i,t} \perp sales\_pc_{i,t}$. As explained above, this requirement should be satisfied because an increase in population should not lead to more books being bought (or read) per person.

Given that exact information on total population are not available on a monthly basis on the federal state level, we use labor force as a proxy for population.[25] The labor force comprises persons that, according to the definition of the OECD, "fulfil the requirements for inclusion among the employed (civilian employment plus the armed forces) or the unemployed".[26] Thus, labor force basically captures the share of the population available to

[23]Note that, here, "*ceteris paribus*" includes, in particular, population density and other socio-demographic factors such as education. As explained in Section 4.2, fixed effects are used to control for these factors. See also Section 5.

[24]See Appendix B.2 for the first stage results of our IV approach.

[25]The data was obtained from the German Federal Statistical Office's database (https://bit.ly/2OZ5w6n).

[26]https://data.oecd.org/emp/labour-force.htm (last accessed October 12, 2022).

the labor market, i.e., children, elderly and disabled persons are excluded.[27]

4.2 Regression Analyses

This section presents the results of the analysis for the period 2011−2017 and focuses on sales of physical books. Before we present the regression results, we first explain our 2SLS estimation approach. The structural equation of our basic model takes the following form:

$$sales\_pc_{i,t} = \beta_1 \# \widehat{stores}_{i,t} + \beta_2 gtrends_{i,t} + \beta_3 ereader_{i,t} + \xi_i + \xi_t + u_{i,t}, \quad (1)$$

where the dependent variable are the sales of print books per capita in federal state i and month t. Our treatment variable $\#\widehat{stores}_{i,t}$ refers to the fitted values from the first-stage for the absolute number of bookstores in federal state i in month t.

The year-month fixed effects are given by ξ_t. As mentioned above, book sales are decreasing over time. This trend, which could, for instance, be associated with more consumers substituting reading with movie streaming or online gaming, is captured by those dummies. Moreover, the year-month fixed effects also control for the seasonality of book sales, e.g., the large Christmas effect. The variable $gtrends_{i,t}$ captures the Google Trends Index for the topic "book", $ereader_{i,t}$ gives the Google Trends Index for the search item "E-book reader" and ξ_i depicts time-invariant fixed effects for each

[27]On a yearly basis, the Pearson correlation coefficient between the labor force and the population size is 0.9672, i.e., population and labor force are strongly correlated. The measure is thus a close proxy to total population.

federal state i.[28]

As explained in Section 4.1, we use $pop_{i,t}$ and $pop_{i,t}^2$ as instruments for the potentially endogenous variable $\#stores_{i,t}$. $\#stores_{i,t}$ is correlated with $pop_{i,t}$ and $pop_{i,t}^2$, which is a prerequisite for the relevance condition. Note that, for the relevance condition to be satisfied, $pop_{i,t}^2$ needs to be implemented, as otherwise the correlation is not sufficiently strong. As explained in Section 4.1, the orthogonality assumption should be satisfied as well because it is reasonable to assume that population does not affect how many books a person reads. Thus, the projection in the first stage regression of our base model can be formalized as follows:

$$\#stores_{i,t} = \delta_1 pop_{i,t} + \delta_2 pop_{i,t}^2 + \delta_3 gtrends_{i,t} + \delta_4 ereader_{i,t} + \xi_i + \xi_t + \varepsilon_{i,t}. \quad (2)$$

The corresponding first stage regression results are presented in Table B.2 of Appendix B.2.

4.2.1 Baseline Results

Table 1 presents the regression results for Equation (1).[29] We run the regression for total sales, i. e. the sum of sales in physical bookstores and

[28]Note that we only investigate 14 out of 16 federal states in Germany. We excluded two federal states (Saarland and Thuringia) from our analyses due to potential errors in the data.

[29]We present the results of a naive OLS estimation for Equation (1) in Table B.1 of Appendix B.1. In this regression, the number of bookstores has a positive effect on the print book sales, which is significant on the 5% level. However, as discussed in the previous sections (esp. Section 4.1), this naive OLS regression is prone to endogeneity. Neglecting this endogeneity issue could lead to a biased estimation, which is the reason why we apply an IV estimation approach.

e-Commerce (column (1)), and for the two different sales channels physical bookstores and e-Commerce separately (columns (2) and (3)).

Table 1: Main estimations with sales volumes on print books per capita as dependent variable.

	(1) Aggregated	(2) Offline	(3) e-Commerce
# Stores	0.000248***	0.000183***	0.0000559*
	(0.0000611)	(0.0000486)	(0.0000260)
Google Trends	0.00376***	0.00317***	0.000739*
	(0.000811)	(0.000637)	(0.000354)
Google Trends e-Reader	0.000470	0.000169	0.000225
	(0.00113)	(0.000815)	(0.000622)
Federal State FE	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes
Anderson-Rubin Wald F-statistic	8.009	7.281	2.095
Kleibergen-Paap rk Wald F-statistic	2238.1	2238.1	2238.1
# of observations	1176	1176	1176

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient of $\#stores$ measures how a change in the total number of bookshops affects book sales per capita in an average month for an average federal state. Thus, based on column (1) in Table 1, we can conclude that if one physical bookstore closes, *ceteris paribus*, sales per capita decrease by 0.000248 on average (average month, average federal state). The effect can be disentangled into the offline and online channel. The effect on sales in physical bookstores is 0.000183 and in e-Commerce 0.0000559.

Before turning to the interpretation of the findings, some technical aspects have to be discussed. The potentially endogenous variable in Equation (1), $\#stores$, is identified by the instrument as indicated by the Kleibergen-Paap rk-Wald F-statistic which clearly exceeds the IV critical value from Stock and Yogo (2005). Furthermore, the Null of the Anderson-Rubin Wald F-statistic can be rejected for the regression with aggregated (column (1) in Table 1) and offline sales (column (2)), indicating that the potentially endogenous regressor is relevant in the structural equation. Additionally, a test for exogeneity of the variable *#stores* using the difference of two Sargan-Hansen statistics (also *GMM distance* or *C-statistic*) can be rejected.[30]

A positive coefficient for *#stores* (column (1) in Table 1) means that when a physical bookstore exits the market, aggregate sales decrease. The same applies to sales in the offline and the online channel in isolation (columns (2) and (3) in Table 1). This implies that, on average, market exit of physical bookstores is not fully compensated for by sales in other physical bookstores. (Full compensation would occur if the effects were not statistically significant from zero. In that case, depending on which effect is considered, the closure of a bookstore would not affect total sale, sales in the offline or sales in the online channel.) Our results thus indicate that, on average, market exit of physical bookstores leads to a significant decrease in sales.

The effect is different for the offline and the online channel. One can see that offline sales decrease more strongly upon market exit of physical bookstores than online sales (columns (2) and (3) in Table 1).

The positive coefficient of e-Commerce might seem confusing at first glance. As explained above, e-Commerce and physical bookstores can be considered substitutes such that e-Commerce should become more attractive when physical bookstores close. However, there appears to be complementarity between the channels, as substitutability would require a significantly

[30] This endogeneity test statistic is numerically equivalent to a Durbin-Wu Hausman test under conditional homoskedasticity (see Hayashi (2000)).

negative coefficient of $\#stores$. Even though the exact mechanism cannot be clearly identified based on our data, this observation is consistent with the well-known problem of free-riding on service provision as in, e.g., Telser (1960). The features or "services" of one sales channel affect sales in the other channel. A "showrooming"-effect (Zhang et al., 2018) might be prevalent. Consumers' attention is drawn to a certain title in physical bookstores. This initial impetus leads to higher online sales through online reviews or word-of-mouth (Beck, 2007). Analyzing such a complementarity in more detail is left for future research.

To get a better understanding of the magnitude of our findings, it is illuminating to express the decrease in sales in absolute terms. Population, which we approximate using labor force (see above), is, on average, 3 million per federal state (42 million total). Thus, when one bookstore closes, average book sales decrease by $0.000248 \cdot 3$ million $= 744$ units per month in an average federal state.

Based on these figures, it is possible to compute a counterfactual scenario. This scenario is used to answer the following question: How did the closures of bookstores affect print book sales in the period 2011-2017? In order to answer this question, we will assume that the number of bookstores remains the same as in 2011 for the entire observation period. First, note that, relative to 2011, the total loss in print book sales accumulated to around 136.69 million units from 2012-2017 (see Figure 4). (That is, in 2012 book sales decreased by 8.76 million units relative to 2011; in 2013 an additional 18.58 million units were sold less, such that losses accumulated to around 27.3 million units relative to 2011, and so on.) Next, recall that our estimations predict that the closure of a bookstore leads to an average drop in print book sales of 744 units per month, for an "average" federal state. Note that this reflects monthly sales, i.e., for each bookstore that exits the market, sales decrease by that amount in each subsequent month. For instance, if a bookstore closes in $t = 0$, the loss in sales triggered by that exit amounts to $7 \times 12 \times 744 = 62,496$ over the period 2011-2017. In the 14 federal states examined in this section, 1,382 bookstores exited the market.[31] Given this drop in the number of bookstores, the mean loss in sales triggered by the closure of bookstores sums up to $(62,496 \times 1,382)/2 = 43,184,736$. Thus, the loss in sales triggered by exit of physical bookstores accounts for $43,184,736/136,690,000 \approx 31.59\%$ of the loss in total book sales in the period 2011-2017.

A second, more speculative interpretation starts from the number of books a physical bookstore sells on average per month. These average monthly sales are around 3,760 units. If a physical bookstore exits, these sales will in part go to other physical bookstores and be made online, respectively. In part, they will be lost. In addition, there is the loss due to the showrooming effect. This effect amounts to 0.0000559×3 million ≈ 168 (see column (3) in Table 1). Therefore, notional sales of the exiting bookstore were about 3,927. As the aggregate loss of books is 744, we know that about $744/3,927 = 19\%$ of sales are lost. If we note that around 20% of bookstores were closed from 2011-2017, we obtain a result, which is in the same order of magnitude as the above one, if we apply the following simple rule of thumb. With 20% of bookstores exiting and an aggregate sales decrease by around 19% upon

[31] Including Thurungia and Saarland, 1,435 were closed between 2011-2017, see Section 3.

exit of a bookstore (see above), total sales decrease by around 3.8% due to the closure of bookstores. Based on average annual sales of around 277.67 million, this amounts to 10.55 million books. Given an annual sales decrease by around 33.18 million in the period 2011-2017 (see Figure 4), around 31.8% of the decrease in sales can be traced back to the closure of bookstores.[32]

When interpreting these findings, keep in mind that they occur in the German market, which is characterized by a relatively high density of bookstores. According to our data, in Germany there is approximately one physical bookstore per $16,400$ inhabitants in 2018. For instance, in the USA the number of physical bookstores is much lower. The number of stores that sold books peaked at $12,000$ in 1992 (see Wu (2018)), so that there was approximately one physical bookstore per $21,400$ inhabitants. It remains an open question whether the effect of a closure on sales becomes more pronounced the lower the number of physical stores because with each closure it becomes increasingly more difficult for consumers to find an appropriate substitute offline (e.g., increasing physical traveling distances). It is left open for future research to investigate whether the effect of the closure of physical bookstores is different in markets with fewer or more physical bookstores than Germany.

Finally, note that the coefficient of Google Trends for the term “book” is statistically significant. This means that, *ceteris paribus*, with increasing search volume, sales of books tend to increase. In contrast, the coefficient of Google Trends e-Reader is insignificant in Table 1 implying that the search behavior of the consumers for e-Readers does not significantly affect the

[32]Note when we account for the actual average sales of a bookstore (see above), i.e. 3,760, this figure would be 33.1%.

physical book sales per capita. In other words, it does not appear to be the case that consumers substitute e-Books for print books in a systematic fashion, as would be indicated by a negative coefficient for that index. This can also be seen as an indication that consumers do not substitute print books by e-Books when bookstores close, or at least not one-for-one. E-Book sales will be discussed in Section 4.3 in more detail.

4.2.2 Genre-specific Effects

In this section, we will check whether there are genre-specific differences in the effects of exit of physical bookstores on sales. In doing so, we run the IV regressions explained above for the nine different book genres (see in Section 3) separately. Table 2 represents the results for our treatment variable $\#stores$ in the respective second stages of the 2SLS regressions with aggregated sales, offline sales and sales in e-Commerce as dependent variables.

The results presented in the first column of Table 2 are largely in line with the results presented above. For instance, if one bookstore closes, 0.0000692 print books of the fiction genre are sold less in both retail channels on average (month and federal state). This means that $0.0000692 \cdot 3$ million $= 207.6$ fiction books are sold less per month, on average. Similar interpretations can be applied to the genres non-fiction, humanities, children books, natural sciences, guidebooks and travel. In particular, note that the effect for children books is even slightly stronger than that for fiction ($0.0000701 \cdot 3$ million $=$ 210.3).

However, the change in the number of physical bookstores has no significant impact on offline sales of school books and books belonging to the

Table 2: IV regressions by genres.

	Aggregated	Sales Offline	Sales e-Commerce
fiction			
# Stores	0.0000692***	0.0000557***	0.0000151*
	(0.0000191)	(0.0000166)	(0.00000678)
non-fiction			
# Stores	0.0000150**	0.0000112**	0.00000379
	(0.00000459)	(0.00000359)	(0.00000215)
humanities			
# Stores	0.0000106***	0.00000695***	0.00000209
	(0.00000200)	(0.00000138)	(0.00000125)
children books			
# Stores	0.0000701***	0.0000573***	0.0000113*
	(0.0000109)	(0.00000914)	(0.00000455)
natural sciences			
# Stores	0.00000655***	0.00000358***	0.00000319*
	(0.00000143)	(0.000000474)	(0.00000131)
guidebooks			
# Stores	0.0000400***	0.0000323***	0.00000975
	(0.00000766)	(0.00000413)	(0.00000588)
travel			
# Stores	0.0000128***	0.0000106***	0.00000233*
	(0.00000251)	(0.00000211)	(0.00000118)
school books			
# Stores	0.00000456	0.00000533	0.00000281
	(0.0000369)	(0.0000332)	(0.00000520)
social sciences			
# Stores	0.00000276	0.000000745	0.00000275
	(0.00000184)	(0.000000853)	(0.00000148)

Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

category social sciences. In other words, if physical bookstores close, total sales as well as sales in the offline and online category remain unaffected. This implies closer substitution patterns between the two channels in case of school books. This finding is in line with expectations as the demand for school books can be expected to be highly inelastic. In that genre, service aspects such as expert opinion, ad hoc sales, showrooming, etc. should be almost irrelevant. Thus, on average, there appears to be perfect substitution between sales in different physical bookstores or between physical bookstores and e-Commerce in this category.

The coefficients in the third column in Table 2 (Sales e-Commerce) require some discussion. As already explained, these coefficients capture the effect of a change in the number of bookstores on sales in e-Commerce. A positive coefficient indicates a showrooming-effect. This is the case for fiction, children books, natural sciences and travel. However, note that this effect is only significant on the 5%-level. Moreover, for fiction (45.3 units) and children books (33.9 units) the effect is far stronger than for natural sciences (9.57) and travel (6.99). In any case, these results suggest the existence of a showrooming-effect, which increases e-Commerce sales.

4.2.3 Effects On Sales In Different Channels

Finally, the effects can be broken down by four different sales channels: independent stores, chain stores, station bookshops and e-Commerce.[33] In doing so, the same regressions as presented in Section 4.2 were run separately for each sales channel. We therefore explain the sales in each channel by the total number of bookstores. Note that we cannot identify to which category an exiting or entering physical bookstore belongs. Against this background, the respective effects on sales of print books in each sales channel shown in Table 3 occur when an "average" physical bookstore closes.[34]

The results depicted in Table 3 are largely in line with the previous findings with respective to the general direction of the effects. The results suggest

[33]Note that the results for the e-Commerce channel (column (2)) are the same as in Table 1. They were added for completeness.

[34]For this reason, we cannot identify the substitution patterns between the different types of bookstores in more detail.

Table 3: Estimation by physical bookstore types.

	(1) independent	(2) e-Commerce	(3) retail chain	(4) station bookstores
# stores	0.0000729**	0.0000559*	0.000122**	-0.000000573
	(0.0000239)	(0.0000260)	(0.0000394)	(0.00000448)
Google Trends	0.00115***	0.000739*	0.00158**	0.000175**
	(0.000313)	(0.000354)	(0.000516)	(0.0000535)
Google Trends e-Reader	-0.000297	0.000225	0.000424	-0.0000202
	(0.000541)	(0.000622)	(0.000598)	(0.0000364)
Federal State FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes
Anderson-Rubin Wald F-statistic	4.238	2.095	8.415	47.09
Kleibergen-Paap Wald F-statistic	2238.1	2238.1	2238.1	2238.1
# of observations	1176	1176	1176	1176

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

that upon closure of a bookstore, on average, sales in independent bookstores and retail chains decrease (given the caveats just mentioned). The effect on sales in station bookstores is insignificant. This means that book sales in station bookstores appear not be affected by the number of bookstores. This could be explained by the business model of station bookstores, which should mostly be based on selling books to travellers. This means that the service-dimension described above can be expected to play a minor role in station bookstores.

4.3 Estimation With E-Books

It is reasonable to assume that at least some consumers substitute e-Books for print books when physical bookstores close, especially since there is evidence that e-Books expanded book sales (see Gilbert, 2015, p. 167–170, for an overview). It is thus appropriate to check whether the effects of exit of physical stores on book sales presented above are overestimated due to missing information on e-Book sales. Thus, we will now deploy the consumer panel data set that includes information on e-Book sales (see Appendix A.2 for more details). The data is used to repeat the analyses presented in Section 4.2. Estimation results are reported in Table 4 below.

We run the regression for total book sales (column (1)) and for the two book formats print and e-Books separately (columns (2) and (3)). As shown in Table 4, the effect of the number of physical bookstores on book sales remains statistically significant in all three regression approaches. The reported coefficient in column (1) implies a decrease in total book sales per capita, including e-Books, of 0.00257 in a given quarter per federal state (significant on the 1%-level), when a physical bookstore closes. (For a more detailed discussion of the order of magnitude of the effect, see below.) The effect in the number of bookstores on book sales is also significantly positive when we run the regression for print and e-Books separately (columns (2) and (3)). Thereby, the number of physical bookstores has a larger effect on print book sales per capita (0.00239) than on e-Book sales per capita (0.000189). Again, the significant result for e-Books might be explained by a "showrooming"-effect.

Table 4: Estimation with total book sales (print books and e-Books) per capita as dependent variable using consumer survey data from GfK.

	(1) All books	(2) Print	(3) e-Books
# Stores	0.00257***	0.00239***	0.000189***
	(0.0000841)	(0.0000849)	(0.0000228)
Federal State FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Anderson-Rubin Wald F-statistic	479.6	443.6	29.27
Kleibergen-Paap rk Wald F-statistic	8475.0	8475.0	8475.0
# of observations	320	320	320

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

These findings also show that the main results reported in Section 4.2 are qualitatively robust when information on e-Book sales are included into the analyses, as e-Book sales apparently do not compensate for the decrease in sales triggered by the exit of physical bookstores, as this would require a negative coefficient of $\#stores$. Thus, a significantly positive (or insignificant) coefficient does not challenge the previous findings.

Note that the effects appear to be stronger when we use consumer panel data. The coefficient for "All books" reported in column 1 implies a decrease in average quarterly sales of $0.00257 \cdot 3$ million $= 7,710$, i.e., 2,570 units per month. This effect is around 3.5-times stronger than the one reported in the Section 4.2.1. A possible explanation for this difference is that information in the consumer panel data set are extrapolated by GfK based on survey data, whereas the scanner data are based on actual sales (see Section 3). Thus, the two data sets differ in terms of absolute sales figures. As explained in Section 3, we consider the scanner data to be a more accurate description of actual book sales than survey data.

5 Robustness Checks

Robustness checks were performed for the analyses presented in the previous sections. All estimation results are presented in Appendix C. Our results remain robust with respect to the following modifications.

- To rule out that book prices affect some of the results, the change in the monthly average book prices is integrated as a "bad" control variable into the second stage of the IV-regression. Average book prices are computed based on scanner data for the period $2011-2017$. Including prices as a control can rule out that decreasing book sales are driven by increasing book prices and not by the decreasing number of physical bookstores. The control variable is "bad" because the relationship between sales and prices is bi-directional (see Appendix C.1).

- The number of students is included as a covariate. This is done to control for the education levels in the 14 German federal states of our data set. The fixed effects included into the regressions capture nation-wide time trends and time-invariant effects that affect individual federal states. However, socio-demographic effects such as education can be time-variant and specific to a federal state. Those effects potentially confound our analyses (see Appendix C.2).

- The analyses presented in the Section 4.2 can be performed using a combination of survey and scanner data to include e-Book sales. Therefore, we use the ratio of e-Book sales from the survey data to calculate absolute e-Book sales based on the scanner data. This e-Book sales data then can be used to repeat our estimations for the period $2014-2017$ (quarterly level) with data on print and digital book sales (see Appendix C.3).

- We also regress the book sales per capita on the lagged number of physical bookstores in an OLS estimation approach. Lagging our treatment variable $\#stores$ does at least partially solve the reverse causality issue between the number of bookstores and the book sales per capita (see Appendix C.4).

6 Conclusion

We find that, overall, e-Commerce does not pose a perfect substitute to physical retailers from the viewpoint of the consumers. Our results predict that when a physical bookstore closes, on average, monthly book sales decrease by 744 units. From 2011–2017, 1,382 bookstores were closed across the 14 federal states comprising our data set. The loss in sales thus accumulates to, on average, around 43 million units. In the same period, the loss in total print book sales accumulates to around 137 million units. Therefore, our results indicate that around one third of the drop in print book sales can be traced by to the closures of bookstores. Apparently, a large enough number of consumers prefers to buy books at physical bookstores for closures of those stores to have a statistically significant, negative impact on the total sales of books. Consumers' preferences towards offline "services" might potentially be affected by expert opinion, a more careful selection and presentation of titles, ad-hoc purchases or simply the atmosphere of physical bookstores.

The magnitude of the effect differs between genres. The effect is particularly strong for fiction and children books titles, where we not only find a significant drop in offline sales but also in *online* sales following the market exit of physical bookstores. A potential interpretation would be a showrooming-effect, which indicates complementarity between the sales channels when it comes to fiction titles. On the other hand, we find a high degree of substitutability between the two channels when it comes to school books. This observation appears intuitive as the demand for schoolbooks should be determined by reasons other than features or services offered by online and offline bookstores.

It is important to note that the magnitude of the effects of market exit of physical bookstores was determined for the German market, which has a relatively high number of physical bookstores. It remains an open question whether the effect of a closure on sales becomes more pronounced the lower the number of physical stores because with each closure it becomes increasingly more difficult for consumers to find an appropriate substitute offline (e.g., increasing physical traveling distances). It is left open for future research to investigate whether the effect of the closure of physical bookstores is different in markets with fewer or more physical bookstores than Germany.

Our finding that consumers perceive online and offline retailers as imperfect substitutes provides an additional facet to the policy evaluation of fixed book prices. The goal of these vertical restraints is usually to protect books as cultural or merit goods. It is described in the literature that fixed book prices can promote market entry (Bouckaert (2000), Guo and Lai (2017), Elzinga and Mills (2008, p. 1848)) and prevent exit by, e.g., securing margins, in particular in the presence of online competition (Marvel and McCafferty (1985, p. 376),Bouckaert (2000), Guo and Lai (2017), Elzinga and Mills (2008, p. 1848), Legros and Stahl (2019)). There is evidence from the UK that after the abrogation of the Net Book Agreement, UK's and Ireland's fixed book price system abandoned in the 1990's, the book market consolidated (Davies et al., 2004; Dearnley & Feather, 2002; Fishwick et al., 1997). In combination with our finding that a larger number of physical bookstores promotes book sales, fixed book price systems may thus support the policy goal of securing a broad supply of books. Even though systematic analysis is warranted, this efficiency effect of fixed book prices would have to be taken into account when evaluating the welfare effects of suppressing price competition among retailers.

References

Akman, P., & Sokol, D. D. (2017). Online rpm and mfn under antitrust law and economics. *Review of Industrial Organization*, *50*(2), 133–151.

Bagwell, K., & Riordan, M. H. (1991). High and declining prices signal product quality. *The American Economic Review*, 224–239.

Balasubramanian, S. (1998). Mail versus mall: A strategic analysis of competition between direct marketers and conventional retailers. *Marketing Science*, *17*(3), 181–195.

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.

Bell, D. R., Gallino, S., & Moreno, A. (2017). Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, *64*(4), 1629–1651.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

Bester, H. (1998). Quality uncertainty mitigates product differentiation. *The RAND Journal of Economics*, 828–844.

Bouckaert, J. (2000). Monopolistic competition with a mail order business. *Economics Letters*, *66*(3), 303–310.

Brown, J. R., & Goolsbee, A. (2002). Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of political economy*, *110*(3), 481–507.

Brynjolfsson, E., Hu, Y., & Rahman, M. S. (2009). Battle of the retail channels: How product selection and geography drive cross-channel competition. *Management Science*, *55*(11), 1755–1765.

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

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

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless commerce? a comparison of internet and conventional retailers. *Management science*, *46*(4), 563–585.

Burt, S., & Sparks, L. (2003). E-commerce and the retail process: A review. *Journal of Retailing and Consumer Services*, *10*(5), 275–286.

Chen, H., Hu, Y. J., & Smith, M. D. (2019). The impact of e-book distribution on print sales: Analysis of a natural experiment. *Management Science*, *65*(1), 19–31.

Chen, P.-Y., Wu, S.-y., & Yoon, J. (2004). The impact of online recommendations and consumer feedback on sales. *ICIS 2004 Proceedings*, 58.

Chevalier, J., & Goolsbee, A. (2003). Measuring prices and price competition online: Amazon. com and barnesandnoble. com. *Quantitative marketing and Economics*, *1*(2), 203–222.

Chu, C.-P., Guo, W.-C., & Lai, F.-C. (2012). On the competition between an online bookstore and a physical bookstore. *NETNOMICS: Economic Research and Electronic Networking*, *13*(3), 141–154.

Clay, K., Krishnan, R., Wolff, E., & Fernandes, D. (2002). Retail strategies on the web: Price and non–price competition in the online book industry. *The Journal of Industrial Economics*, *50*(3), 351–367.

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.

Couture, V., Faber, B., Gu, Y., & Liu, L. (2021). Connecting the countryside via e-commerce: Evidence from china. *American Economic Review: Insights*, *3*(1), 35–50.

Cowen, T. (2008). Why everything has changed: The recent revolution in cultural economics. *Journal of Cultural Economics*, *32*, 261–273.

Davies, S., Coles, H., Olczak, M., Pike, C., & Wilson, C. (2004). The benefits from competition: Some illustrative uk cases. *DTI Economics Papers*, *9*.

Dearnley, J., & Feather, J. (2002). The UK Bookselling Trade Without Resale Price Maintenance: An Overview of Change 1995–2001. *Publishing Research Quarterly*, *17*(4), 16–31.

Elzinga, K. G., & Mills, D. E. (2008). The economics of resale price maintenance. *Issues in Competition Law and Policy (3-Volume Set), Kenneth G. Elzinga & David E. Mills, eds., ABA Section of Antitrust Law.*

Fishwick, F., Fitzsimons, S., & Muir, L. H. (1997). *Effects of the Abandonment of the Net Book Agreement: First Interim Research Report*. Book Trust London.

Gauri, D. K., Jindal, R. P., Ratchford, B., Fox, E., Bhatnagar, A., Pandey, A., Navallo, J. R., Fogarty, J., Carr, S., & Howerton, E. (2021). Evolution of retail formats: Past, present, and future. *Journal of Retailing*, *97*(1), 42–61.

Gilbert, R. J. (2015). E-books: A tale of digital disruption. *Journal of Economic Perspectives*, *29*(3), 165–184.

Goldmanis, M., Hortaçsu, A., Syverson, C., & Emre, Ö. (2010). E-commerce and the market structure of retail industries. *The Economic Journal*, *120*(545), 651–682.

Goolsbee, A. (2000). In a world without borders: The impact of taxes on internet commerce. *The Quarterly Journal of Economics*, *115*(2), 561–576.

Goolsbee, A. (2001). Competition in the computer industry: Online versus retail. *The Journal of Industrial Economics*, *49*(4), 487–499.

Guo, W.-C., & Lai, F.-C. (2017). Prices, locations and welfare when an online retailer competes with heterogeneous brick-and-mortar retailers. *The Journal of Industrial Economics*, *65*(2), 439–468.

Gürtzgen, N., Diegmann, A., Pohlan, L., & van den Berg, G. J. (2021). Do digital information technologies help unemployed job seekers find a job? evidence from the broadband internet expansion in germany. *European Economic Review*, *132*, 103657.

Hayashi, F. (2000). Econometrics,—— princeton university press: Princeton.

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.

Jin, G. Z., & Kato, A. (2007). Dividing online and offline: A case study. *The review of economic studies*, *74*(3), 981–1004.

Legros, P., & Stahl, K. O. (2019). Local search markets and external competition.

Li, H. (2015). Cannibalization or market expansion? the impact of e-books on print book sales. *The Impact of E-Books on Print Book Sales (February 20, 2015).*

Li, H. (2019). Intertemporal price discrimination with complementary products: E-books and e-readers. *Management Science*, *65*(6), 2665–2694.

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics*, 214–231.

Loginova, O. (2009). Real and virtual competition. *The Journal of Industrial Economics*, *57*(2), 319–342.

Mantovani, A., Piga, C. A., & Reggiani, C. (2021). Online platform price parity clauses: Evidence from the eu booking. com case. *European Economic Review*, *131*, 103625.

Marvel, H. P., & McCafferty, S. (1985). The welfare effects of resale price maintenance. *The Journal of Law and Economics*, *28*(2), 363–379.

Marvel, H. P., & McCafferty, S. (1990). Resale price maintenance and quality certification. *J. Reprints Antitrust L. & Econ.*, *20*, 251.

Mérel, P., Ortiz-Bobea, A., & Paroissien, E. (2021). How big is the "lemons" problem? historical evidence from french wines. *European Economic Review*, *138*, 103824.

Nelson, P. (1974). Advertising as information. *Journal of Political Economy*, *82*(4), 729–754.

Ofek, E., Katona, Z., & Sarvary, M. (2011). "bricks and clicks": The impact of product returns on the strategies of multichannel retailers. *Marketing Science*, *30*(1), 42–60.

Poort, J., & van Eijk, N. (2017). Digital fixation: The law and economics of a fixed e-book price. *International Journal of Cultural Policy*, *23*(4), 464–481.

Prince, J. T. (2007). The beginning of online/retail competition and its origins: An application to personal computers. *International Journal of Industrial Organization*, *25*(1), 139–156.

Rabinovich, E., Sinha, R., & Laseter, T. (2011). Unlimited shelf space in internet supply chains: Treasure trove or wasteland? *Journal of Operations Management*, *29*(4), 305–317.

Reimers, I. (2016). Can private copyright protection be effective? evidence from book publishing. *The journal of law and economics*, *59*(2), 411–440.

Reimers, I., & Waldfogel, J. (2017). Throwing the books at them: Amazon's puzzling long run pricing strategy. *Southern Economic Journal*, *83*(4), 869–885.

Reimers, I., & Waldfogel, J. (2021). Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, *111*(6), 1944–71.

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.

Riordan, M. H. (1986). Monopolistic competition with experience goods. *The Quarterly Journal of Economics*, *101*(2), 265–279.

Schulz, N., & Stahl, K. (1996). Do consumers search for the highest price? oligopoly equilibrium and monopoly optimum in differentiated-products markets. *The RAND Journal of Economics*, 542–562.

Sinai, T., & Waldfogel, J. (2004). Geography and the internet: Is the internet a substitute or a complement for cities? *Journal of Urban Economics*, *56*(1), 1–24.

Somogyi, R. (2020). Bertrand–edgeworth competition with substantial horizontal product differentiation. *Mathematical Social Sciences*, *108*, 27–37.

Srinivasan, S. S., Anderson, R., & Ponnavolu, K. (2002). Customer loyalty in e-commerce: An exploration of its antecedents and consequences. *Journal of retailing*, *78*(1), 41–50.

Stahl, K. (1982). Differentiated products, consumer search, and locational oligopoly. *The Journal of Industrial Economics*, 97–113.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression, in dwk andrews and jh stock, eds., identification and inference for econometric models: Essays in honor of thomas j. rothenberg. cambridge: Cambridge university press.

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.

Telser, L. G. (1960). Why should manufacturers want fair trade? *The journal of law and economics*, *3*, 86–105.

Waldfogel, J. (2017). How digitization has created a golden age of music, movies, books, and television. *Journal of economic perspectives*, *31*(3), 195–214.

Waldfogel, J. (2018). Digital renaissance. In *Digital renaissance*. Princeton University Press.

Waldfogel, J., & Reimers, I. (2015). Storming the gatekeepers: Digital disintermediation in the market for books. *Information economics and policy*, *31*, 47–58.

Wang, K., & Goldfarb, A. (2017). Can offline stores drive online sales? *Journal of Marketing Research*, *54*(5), 706–719.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Wu, S. X. (2018). Adaptation or death? bookstore chains meet online competition. *Economic Research Initiatives at Duke (ERID) Working Paper*, (273), 18–70.

Zhang, T., Ge, L., Gou, Q., & Chen, L. (2018). Consumer showrooming, the sunk cost effect and online-offline competition. *Journal of Electronic Commerce Research*, *19*(1), 55–74.

Appendix

A Data structure

A.1 Scanner data

Table A.1: Characteristics of the scanner data

Data Characteristics	Description
Data Provider	media control GmbH
Data Collection Method	Store data (POS)
Measurements	Revenue, Titles Sold
Coverage	2011–2017
Frequency	Monthly
Geographical Scope	Federal State
Granularity	Sales channel & genre
Sales channels	Independent bookstores, chain stores, e-commerce, drug stores, electronics stores, department stores, railway stations, grocery stores
Genres	Fiction, non-fiction, humanities, children books, natural sciences, guidebooks, travel, school books, social sciences
Formats	Print (hard cover, soft cover/paperback)

A.2 Survey data

Table A.2: Characteristics of the survey data

Data Characteristics	Description
Data Provider	GfK GmbH
Data Collection Method	Panel of 20,000 respondents
Measurements	Revenue, Titles Sold
Coverage	2014–2017
Frequency	Quarterly
Geographical Scope	Federal State
Granularity	Format
Formats	Books physical, e-Books, audio books physical, audio books digital

Figure A.1 shows yearly aggregates by format as well as revenue and quantity-based market shares. For illustration of the development, information for the years 2011 – 2013 has been added for total sales and e-Books. The data are only available as yearly aggregates and for the whole of Germany, rather than on the federal state level. Yearly data for e-Book sales and total sales (including print, e-Books and audio books) can be found on Statista.[A1]

Figure A.2 shows quarterly aggregates (2014 – 2017) of revenues by format. Quarterly aggregates (2014 – 2017) of the number of sold copies by format are shown in Figure A.3.

[A1] For e-Book sales see https://de.statista.com/statistik/daten/studie/232191/umfrage/absatz-von-e-books-in-deutschland; last accessed December 20, 2023. For total sales (including e-Books and audio books) see https://de.statista.com/statistik/daten/studie/416380/umfrage/absatz-von-buechern-in-deutschland; last accessed December 20, 2023.

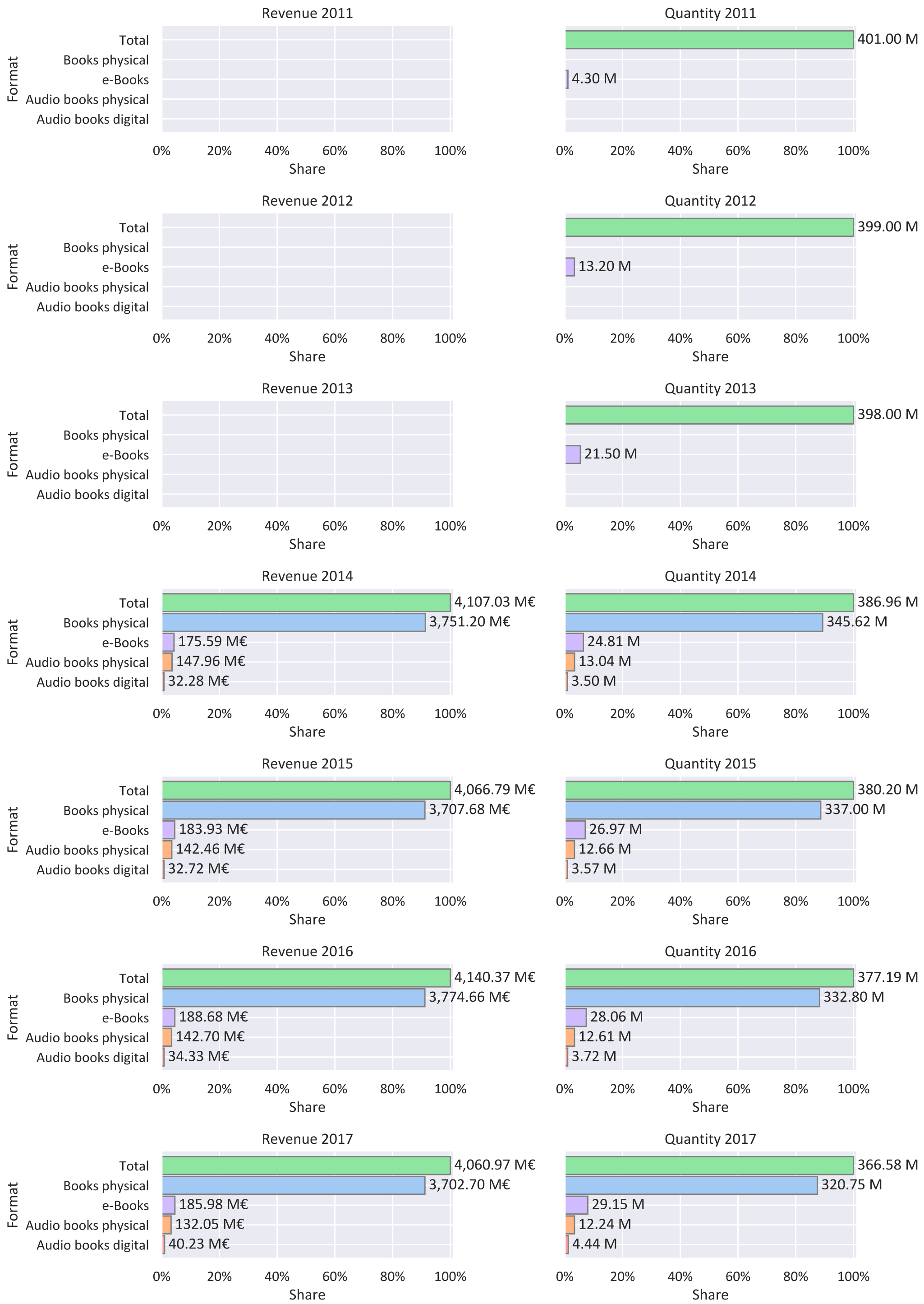

Figure A.1: Evolution of revenue and quantity by format over time (year). Source: GfK GmbH.

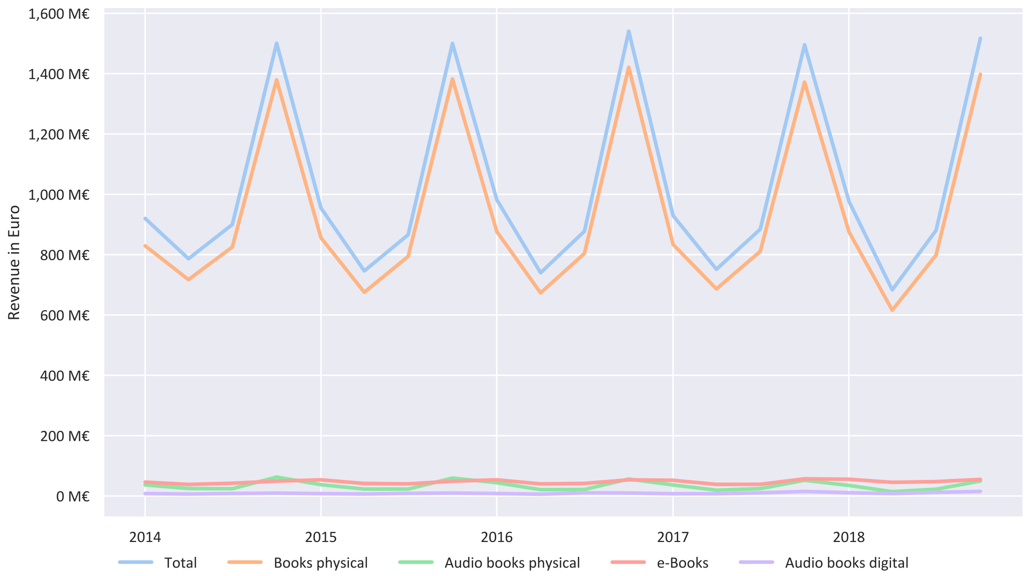

Figure A.2: Evolution of revenues by format over time (quarter). Source: GfK GmbH.

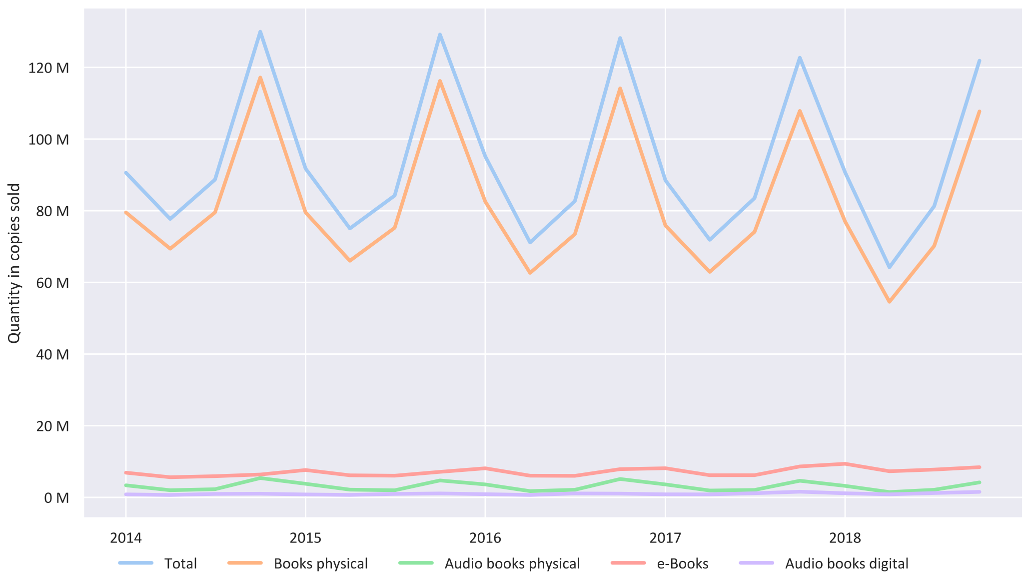

Figure A.3: Evolution of sales volumes by format over time (quarter). Source: GfK GmbH.

B IV–Approach

B.1 Naive OLS estimation of Equation (1)

Table B.1: Naive OLS estimation.

	(1) Aggregated
# Stores	0.000576*
	(0.000259)
Google Trends	0.00341*
	(0.00144)
Google Trends e-Reader	0.000000498
	(0.00117)
Federal State FE	Yes
Year-Month FE	Yes
# of observations	1176

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.2 First stage IV regression result of Equation (2)

Table B.2: First stage regression results of baseline IV estimation.

	(1) # Stores
Pop.	0.000410***
	(0.00000652)
Pop. squared	-4.71e-11***
	(9.16e-13)
Google Trends	0.592**
	(0.184)
Google Trends e-Reader	0.691***
	(0.159)
Federal State FE	Yes
Year-Month FE	Yes
# of observations	1176

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C Robustness checks

C.1 Book prices as "bad" control

The change in the monthly average book prices is integrated as a "bad" control variable into the second stage of the IV-regression. We calculate the change of the monthly average book price for the period $2011 - 2017$ based on our scanner data and use this parameter as a "bad" control variable in our IV estimation approach from Equation (1). Of course, the price is a "bad" control variable in this approach because we now regress a quantity parameter on a price parameter in the second stage of our 2SLS estimation. The results of our IV estimation when controlling for price changes is presented in Table C.1.

Table C.1: IV estimation with monthly average book price change as *bad* control variable.

	(1) Aggregated
# Stores	0.000319***
	(0.0000683)
Google Trends	0.00419***
	(0.00107)
Google Trends e-Reader	0.000595
	(0.00112)
Book Price	-0.00887
	(0.00638)
Federal State FE	Yes
Year-Month FE	Yes
Anderson-Rubin Wald F-statistic	10.24
Kleibergen-Paap Wald F-statistic	1194.4
# of observations	1176

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.1 shows that the newly included variable Book price has no significant effect on the dependent variable total print book sales. The coefficient of this variable is expected to be biased due to reverse causality because quantities are basically regressed on prices. Nevertheless, from Table C.1 one can see that the number of bookstores is still significant and has a positive sign. This implies that a decreasing number of physical bookstores also leads to lower print book sales per capita when we control for book price changes, which is in line with the findings presented in the main text.

C.2 Number of university students as control

State and time specific socio-demographic effects might confound our baseline analysis, in particular changes in the education level. To control for those effects, we use data on the number of university students by federal states[C1] as a proxy for education. Given that this data is only available annually, we aggregate sales taken from scanner data to that level and run a IV regression on a yearly basis. In particular, we regress print book sales per capita on our treatment variable, the number of physical bookstores, and the lagged number of students (as well as other covariates mentioned in Equation (1)) in the second stage of our 2SLS estimation. The results are presented in Table C.2.

Table C.2: IV estimation on a yearly basis when controlling for the number of university students.

	(1) Aggregated
# Stores	0.00642*
	(0.00316)
Google Trends	-0.0351
	(0.0550)
Google Trends e-Reader	0.0828
	(0.0470)
$Students_{t-1}$	-0.000000970
	(0.00000912)
Federal State FE	Yes
Year FE	Yes
Anderson-Rubin Wald F-statistic	2.853
Kleibergen-Paap Wald F-statistic	22.99
# of observations	84

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[C1] This data is provided by the German Federal Statistical Office, see https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000113.

One can see that the lagged number of university students has no significant effect on print book sales per capita in Germany. Note that there is little variation in the number of students such that the federal state fixed effect and the number of students are closely related. However, the lagged number of physical bookstores still has a positive and significant effect on print book sales, which again is in line with the findings presented in the main text. It appears surprising that the result remains statistically significant even though the number of observation is only one twelfth of that in the main text (yearly data instead of monthly data).

C.3 Estimations using a combination of scanner and survey data

In order to incorporate e-Book sales into the scanner data set, we follow these steps:

1. Aggregate the scanner data based on federal state quarters for the period $2014Q1 - 2017Q4$.
2. Utilize survey data to calculate the ratio of e-Book to print book sales for each federal state quarter.
3. Apply the calculated ratio to estimate e-Book sales using the scanner data.

Subsequently, the obtained e-Book sales data can be used to replicate our estimations in Section 4.2 with comprehensive data on both print and digital book sales.

Estimation results using print and calculated e-Book sales for the sales channel e-Commerce as dependent variable are presented in Table C.3. One

Table C.3: IV regression results including e-Book sales.

	(1)
	Sales e-Commerce (incl. e-Books)
# Stores	0.000547***
	(0.0000502)
Federal State FE	Yes
Year-Quarter FE	Yes
Anderson-Rubin Wald F-statistic	60.75
Kleibergen-Paap Wald F-statistic	13927.8
# of observations	224

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

can see that the number of physical bookstores still has a significant positive effect on the book sales per capita in the e-Commerce when also including e-Book sales into the regression for the book scanner data (cf. column (3) in Table 1 of Section 4.2.1). As explained above, this finding can be seen as evidence of a showrooming-effect.

This finding also shows that the findings reported in Section 4.2 are robust when information on e-Book sales are included into the analyses, as e-Book sales apparently do not compensate for the decrease in sales triggered by the exit of physical bookstores, as this would require a negative coefficient of $\#stores$. Thus, a significantly positive (or insignificant) coefficient does not challenge the previous findings.

C.4 OLS Estimation with lagged number of bookstores

We perform an OLS regression of Equation 2 using the monthly lagged number of bookstores as a treatment variable. This estimation strategy should at least partially solve our reserve causality issue between the number of bookstores and the book sales per capita since the book sales per capita in $t = 0$

should not have an effect on the number of bookstores in $t = -1$.

The results of this approach are depicted in Table C.4 and imply that the number of bookstores have a positive effect on the sales per capita in the offline sales channel (column (2)). This means that a closing bookstore in $t = -1$ significantly lowers the book sales per capita in $t = 0$ (significant on the 5% level). However, the effect of a closing bookstore on the book sales per capita is not significant for the aggregated sales and the sales channel e-Commerce (columns (1) and (3)) using this estimation approach.

Table C.4: OLS Estimation with lagged number of bookstores as treatment variable

	(1) Aggregated	(2) Offline	(3) e-Commerce
$\#Stores_{t-1}$	0.000557	0.000562*	-0.00000515
	(0.000272)	(0.000222)	(0.0000983)
Google Trends	0.00276	0.00238*	0.000378
	(0.00168)	(0.000920)	(0.00111)
Google Trends e-Reader	-0.000279	-0.000431	0.000152
	(0.00141)	(0.00105)	(0.00109)
Federal State FE	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes
# of observations	1008	1008	1008

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 The Impact of the Agency Model on E-book Prices: Evidence from the UK

Authors: Maximilian Maurice Gail and Phil-Adrian Klotz

Presented at the following conferences:

- 14. RGS Doctoral Conference, Dortmund (virtual)
- 2021 MaCCI Annual Conference, Mannheim (virtual)
- 48. Annual Conference of the European Association for Research in Industrial Economics (EARIE), 2021 (virtual)
- 15. Annual Competition and Regulation Summer School and Conference (CRESSE), Crete (Greece)

The Impact of the Agency Model on E-book Prices: Evidence from the UK

Maximilian Maurice Gail[a], Phil-Adrian Klotz[a,*]

[a] *Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Licher Strasse 62, Giessen, 35394, Germany*

Abstract

This paper empirically analyzes the effect of the widely used agency model on retail prices of e-books sold in the United Kingdom. Using an unique cross-sectional dataset of e-book prices for a large number of book titles across all major publishing houses, we exploit cross-genre and cross-publisher variation to identify the effect of the agency model on e-book prices. Since the genre information is ambiguous and even missing for some titles in our original dataset, we also apply a Latent Dirichlet Allocation (LDA) approach to determine detailed book genres based on the book's descriptions. Using a matching procedure, we find that retail prices for e-books sold under the agency model are approximately 20% cheaper than book titles sold under the wholesale model. Our results are robust to different regression specifications and double machine learning techniques.

Keywords: e-books, agency agreements, vertical restraints, Amazon, propensity score matching
JEL: D12, D22, L42, L81, L82, Z11
Declarations of interest: None
Acknowledgements: We would like to thank participants at the MaCCI Annual Conference 2021, EARIE Annual Conference 2021 and CRESSE Conference 2021. In particular, we thank Georg Götz, Daniel Herold, Jan Thomas Schäfer, Jona Stinner, Xiang Hui, Franco Mariuzzo and Matthew Olczak for helpful comments. The authors alone are responsible for the content.

---

*Corresponding author
Email address: `phil.a.klotz@wirtschaft.uni-giessen.de` (Phil-Adrian Klotz)

1. Introduction

The rise of the internet - and platform markets more specifically - has accelerated the usage of so-called agency arrangements. Thereby, suppliers pay retailers sales royalties to distribute products at prices determined by suppliers. On the contrary, in many traditional retail markets suppliers charge retailers wholesale prices and retailers set final prices to consumers (*wholesale model*). Even though agency arrangements are also used in some of these conventional markets (e.g., newspapers sold at kiosks), this form of a vertical contract is especially prevalent in online markets (e.g., Amazon Marketplace, Apple App Store, eBay Buy It Now). It is also frequently used in the retail market examined in this paper, namely the e-book market.

In general, books are experience goods because readers can ascertain the quality only after reading a given book title (Nelson, 1970; Reimers and Waldfogel, 2021). In some countries such as Germany, France or Japan, book prices are fixed whereas countries without fixed book price (FBP) systems include the UK and the USA. Fixed book prices are a form of resale price maintenance (RPM) where publishers set retail prices and price competition between retailers is restricted or completely eliminated.[1]Mostly the motivation behind the introduction of FBP systems is the assurance for a broad and diverse supply of books, available through a geographically wide network of bookstores.

[1]Presently, 15 OECD countries have a regulation for fixing the prices of printed books. The fixed prices for printed books typically last 18-24 months after a book has been published.

With the advent of e-books, countries with a FBP system for print books had to decide whether to extend existing legislation to e-books. It is questionable whether the same cultural policy arguments and legal considerations apply, in particular because a geographically wide network of bookstores is irrelevant for e-books. Nevertheless, eight OECD countries with fixed prices for printed books also have fixed prices for e-books, while no country is known to have such a RPM regulation for e-books but not for print books (Poort and van Eijk, 2017). However, in many countries without fixed prices for e-books (such as the UK), this digital product is partly sold under the agency model (Gilbert, 2015), which has similar effects like RPM between a manufacturer and a retailer.[2]

In 2010, Apple in co-operation with the six largest publishing houses have been the first who adopted the agency model for e-books in response to Amazon's aggressive pricing strategy to gain market share. In April 2012, the Department of Justice (DOJ) sued Apple and five of the six publishing houses for conspiring to raise e-book prices by using the agency model in conjunction with most-favored nation (MFN) clauses.[3] Three of the pub-

[2]They are economically similar in the sense that the upstream firms control the retail prices. However, a major difference is that for agency pricing the downstream firms individually delegate retail pricing to the upstream firms, whereas under the classical case RPM is imposed at the market level for any given product.

[3] See United States v. Apple Inc., 12 Civ. 2826 (DLC). MFN clauses stipulate that the retail price of a given product set by a publisher through one retailer must be no higher than the retail price set by this publisher through a competing retailer. Hence, a MFN clause guarantees a retailer who prefers a higher commission, if it raises the commission it charges for one publisher, the retail price will remain the same relative to the other retailers. This effect encourages retailers to push for higher fees, which results in higher retail prices (Johnson, 2017).

lishers settled shortly after the antitrust case was filed, while the other two followed later the same year, which meant that the five publishers could not restrict a retailer's ability to set e-book prices for a period of two years.

Empirical evidence on the price effects of RPM and fixed book prices as well as of the agency model is scarce. While systematic empirical evidence on RPM is limited to case studies (MacKay and Smith, 2017; Ippolito, 1991), the only study investigating the empirical effect of the agency model is the one from De los Santos and Wildenbeest (2017). They have used data on e-book prices of bestselling book titles for the years 2012 and 2013 and the Apple case as an exogenous shock to show that the agency model in combination with MFN clauses led to an average increase in prices between 8-18% (depending on the retailer).[4]

The goal of this study is to analyze the price effect of the agency model using a larger and more detailed data set (especially not only incorporating bestselling book titles) and to check whether similar effects also occur in the absence of an alleged conspiracy as in the Apple case. The internet allows consumers access to a larger number of book titles rather than simply the popular ones (bestselling titles) and different authors have shown the importance of those *long tail* titles in markets for creative products (e.g., Aguiar and Waldfogel, 2018; Brynjolfsson et al., 2003). For instance, Brynjolfsson et al. (2003) have estimated that the benefit consumers obtain from

[4] In a second paper, De los Santos et al. (2018) analyze the switch back from wholesale to agency pricing (after a two-year ban on agency pricing following the Apple case) and find qualitatively similar results for the Amazon platform.

access to *long tail* book titles may be as high as $1.03 billion alone in the year 2000.

Our cross-sectional dataset contains prices for 12,001 e-books published on Amazon UK between 2010 and 2020. Using data from Amazon ensures a high market coverage since Amazon accounted for 50 percent of the UK book sales in 2018.[5] We further use publisher- and book genre variation to estimate the effect of the agency model on e-book retail prices.

The results of our propensity score matching design indicate that e-books sold under the agency model on *Amazon.co.uk* are approximately 20% cheaper than digital books sold under the wholesale model. Various robustness checks qualitatively confirm this main finding. This result contradicts the empirical outcome from De los Santos and Wildenbeest (2017)[6], but fits into explanations put forward by the theoretical literature on agency versus wholesale models (Johnson, 2020; Foros et al., 2017; Gaudin and White, 2014).

The rest of the paper is structured as follows. In Section 2, we discuss

[5] See Nielsen (2018), "Books & Consumers - UK Industry Standard Report Q4 2018", p. 13.

[6] Our paper differs from the study of De los Santos and Wildenbeest (2017) in three important ways. First, while De los Santos and Wildenbeest (2017) use the court decision in the Apple Case (see Footnote 3) as an exogenous shock in their approach, our findings do not rely on an alleged conspiracy. Second, we do not only incorporate bestselling book titles into our empirical analysis but also *long tail* book titles. And third, De los Santos and Wildenbeest (2017) cannot measure the "pure" price effect stemming from agency agreements since those were used in conjunction with MFN clauses at that time. Our price effect can be attributed only to the agency agreements because Amazon has settled with the EU Commission in 2017 not to include MFN clauses in respect of any e-book distributed in the EEA for the next five years (See AT.40153 E-book MFNs and related matters (Amazon), Decision dated May 4, 2017. We have scraped the Amazon data in 2020).

the related literature. Section 3.1 describes the data and we present some descriptive statistics in Section 3.2. Section 4 presents our main estimation strategy and results. In Section 5, our robustness checks are outlined. We conclude and outline the contributions of our paper in Section 6.

2. Related Literature

Our article contributes to several strands of literature. First and foremost, it is related to studies which investigate the competitive effects of the agency model. While the empirical literature on the economic effects of the agency model is rather scarce (apart from the two studies by De los Santos and Wildenbeest (2017) and De los Santos et al. (2018)), several recent theoretical papers have analyzed differences in retail prices between the agency and the wholesale model. Lu (2017) uses a bilateral duopoly model with product differentiation to in the upstream- and downstream market to show that the agency model benefits consumers relative to the wholesale model with lower retail prices due to the elimination of double marginalization. Johnson (2020) finds that when publishers set retail prices instead of retailers (agency model), prices may be higher in early periods but lower in later periods since in the wholesale model retailers initially set low prices to lock in consumers, but find it optimal to raise prices once a sufficient number of consumers are locked in.

Another strand of the theoretical literature on agency models assumes that complementary devices are necessary for the enjoyment of the main products (e.g., an e-book reader in the case of e-books). Gaudin and White

(2014) point out that the incentive of a retailer to set high prices is larger when she has monopolistic control over a complementary device, as it was the case in the e-book market when e-books from Amazon could only be read on a Kindle device. In another model-theoretical setup, Abhishek et al. (2016) show that agency selling is more efficient than the wholesale model and leads to lower retail prices, even though retail prices may be higher under the agency model if there are positive externalities from sales of associated products (such as e-readers in the case of e-books).

Foros et al. (2017) show that the agency model is always anti-competitive (leads to higher retail prices) when it is adopted by the platforms on a market-by-market basis. To be more specific, they find that upstream firms (publishers) will set higher retail prices than downstream firms (retailers) would set if they were in control as long as competition is greater among retailers than among publishers. Moreover, they (ibid.) point out that a retailer who sets retail prices independently (wholesale model) benefits when a horizontal rival is restricted by the agency model since the latter creates a price umbrella, which makes it profitable for the independent price-setting retailer to increase prices. MFN clauses induce industry-wide adoption of agency pricing in their model when such adoption would not otherwise have occurred and can thus be seen as anti-competitive. Condorelli et al. (2018) present a theory that makes the decision whether to use agency or wholesale models endogenously in an environment where the retailer has privileged information about the valuations of consumers and show that retailers prefer the agency model.

More generally, our article is also related to the broader literature on RPM. RPM can lead to lower retail prices due to the internalization of vertical externalities such as double marginalization (Spengler, 1950; Tirole, 1988), it can be used to correct for service externalities (Mathewson and Winter, 1984; Perry and Porter, 1986; Winter, 1993) and RPM regimes can also lead to a larger number of brick-and-mortar (B&M) stores compared to regimes with free prices (Dearnley and Feather, 2002; Davies et al., 2004). However, Rey and Stiglitz (1988, 1994) point out that vertical restraints that eliminate intra-brand competition can also be used to mitigate inter-brand competition and then would be anti-competitive.

Finally, our article also contributes to the newer literature on machine learning (ML) and text mining approaches. Varian (2014) and Athey and Imbens (2019) provide an overview of important ML methods. Wang et al. (2019) use the *Learning to Place* ML approach to predict book sales and find that a strong driving factor of book sales across all genres is the publishing house. We will use DML techniques similar to the approach of Knaus (2021), who has used DML in the case of musical practice of children on cognitive skills and school performance. For a broad overview on text mining approaches see Gentzkow et al. (2019). In our article, we will use a latent Dirichlet allocation (LDA) model to determine book genres by analyzing the book descriptions and expert reviews (e.g., Larsen and Thorsrud, 2019).

3. Data

In this section, we present our dataset containing a large number of book titles. We first describe the construction of our dataset in Section 3.1. Descriptive statistics including information on prices, ratings, reviews and the digital size of e-books (in KB) are presented in Section 3.2.

3.1. Data Set Construction

The data generating process is structured as follows. We have scraped the *Amazon.co.uk* webpage for a list of publisher and imprint names starting mid February 2020 taking two weeks to get e-book prices as well as further book characteristics available on the Amazon website. Therefore, we use *a priori* a list of publishing houses, publishers and imprints, which is taken from a historical *Sunday Times* bestseller list. This procedure ensures that our sample only contains books from publishers with a relatively high market share.[7]

This proceeding also incorporates books into our dataset which have been published before 2019 since we have done the publisher search on *Amazon.co.uk* independently of the format. Thus, it may have happened that for a certain book title, which we have found within our observation period, another format of the same title has already been published a few years ago. However, we have ensured that no book is included in our working data set which has been published earlier than 2010.

[7] The used bestseller list contains entries from January, 2006 until the end of March, 2019.

Our raw dataset consists of roughly one million observations, whereby one observation contains several information on different prices, formats, descriptions, ratings, reviews etc. being available on the Amazon website. For every book title there are three entries if all formats (hardcover, paperback, e-book) are available for a certain book title. However, due to the usage of web scraping methods the dataset contains of some entries that are duplicates or not of interest for our analysis. Hence, after the data cleansing process our working dataset consists of $77,629$ observations, respectively $47,161$ unique book titles.[8] We only use book titles in our estimation approach for which all explanatory variables (book characteristics) are available (see Table 1). Thus, for our empirical analysis 12,001 e-book titles remain in the final working dataset.

Our variables of interest are the retail price, which is the price a consumer must pay for a certain e-book, and the treatment variable *Agency*, which takes the value one if there is a text field on the Amazon webpage of a book title expressing *'This price was set by the publisher'* and zero otherwise.[9] Beyond, we have data on several control variables for our empirical analysis. These variables comprise book characteristics as the book format, the book genre, the size of an e-book in KB, variables on book reviews as the star rating and the number of consumer and expert reviews, variables containing information on the author or publisher of a book title and other

[8] In the former number all three format types (hardcover, paperback, and e-book) are included.

[9] See Appendix Figure A.1 for an example from the Amazon webpage.

variables as the publication date or the recommended retail price (RRP). Table 1 summarizes the descriptions for all variables included in our dataset.

Variables	Information
Price	Retail price from the upper right *Buy-Box*
Format	Hardcover, paperback, Kindle
Star rating	Average rating normalized to be between 0 and 1
No. customer reviews	Number of consumer reviews
No. expert reviews	Number of expert reviews on Amazon
Series	Dummy variable whether book is part of a series
Description and reviews	Detailed text-information on the book and by different reviewers
Genre	Constructed by LDA from the descriptions and reviews (see Appendix B)
RRP	Recommended retail price which is the print RRP. For Kindle it is either related to the hardcover or paperback RRP
Agency	Dummy variable to be one if the price was set by the publisher and zero otherwise. Only possible for e-books
Seller	Sample is restricted to be sold by Amazon
Author	Information on the author of a book
Title	Information on the title of a book
Kindle_Size	Kindle file size (in KB)
Publisher	Name of the publisher. We have different levels of aggregation (Imprint,Publisher,Publishing House)
Amazon rank	Uncategorized Amazon bestseller rank for either print books or e-books
Bestsellers	Number of bestsellers in the Sunday Times Bestseller List conditional on the Author's name
WeekInChart	Average number of weeks in the bestseller charts conditional on the Author's name
Identifier	Aggregation of ASINs to verify the books
Date Retail	Period of time since the publication of a book title (in years)

Table 1: Relevant Variables per book title and the information content they provide.

We have also matched the dataset obtained from Amazon with a historical *Sunday Times* bestseller list to identify authors who have already written a bestselling book title in the past. This variable is important for our empirical analysis (in which we estimate the retail price on an e-book) since the name of a bestseller author is an important quality signal for the book readers.

Each book is a unique product written by an author and mostly published by one publisher. Thus, books are heterogeneous goods which makes it impossible to actually compare the value of one specific book with one another. In order to provide an acceptable analysis, it is therefore also necessary to control for the genres of the several books. However, the genre information is ambiguous and even not available for some book titles on the Amazon webpage.[10] Thus, we use a Latent Dirichlet Allocation (LDA) approach to derive book genres from the descriptions and reviews of the individual books available on the Amazon webpage. This control variable should be able to capture specific effects between the individual genres. We describe this text mining approach in Appendix B.

3.2. Descriptive Statistics

Our final sample consists of 47,161 book titles that have been published on *Amazon.co.uk* by the publishers Bloomsbury, Faber, Hachette, Harper-Collins, Oxford, Pan Macmillan, Penguin Random House, Scholastic, Simon & Schuster, and a group of smaller publishers in the period between 2010 and 2020. However, in overall there are 77,629 observations in our dataset since there are several formats available for some book titles. Even though the focus of our empirical analysis is on the price of e-books, in this section we also present some descriptive statistics for the book formats hardcover and paperback to show the relationship between those three book formats.

Appendix Table A.1 offers descriptive statistics on the variables we use for our empirical analysis, summarized by publishers. In addition to e-book retail prices and the RRP, we also observe several characteristics for each book title such as the sales rank of a title on Amazon, the customer ratings, the number of customer and expert reviews, and the number of pages. As shown in Table A.1, e-books from Scholastic exhibit the lowest

[10] In Section 4.2, we also present one specification with the Amazon genre information as a control variable.

average retail price, while the e-books from Bloomsbury have the highest mean prices. Beyond, the titles from Hachette have the lowest average book rank and the book titles published by Simon & Schuster exhibit the highest average number of customer reviews. Most of the other book characteristics are very similar across publishers.

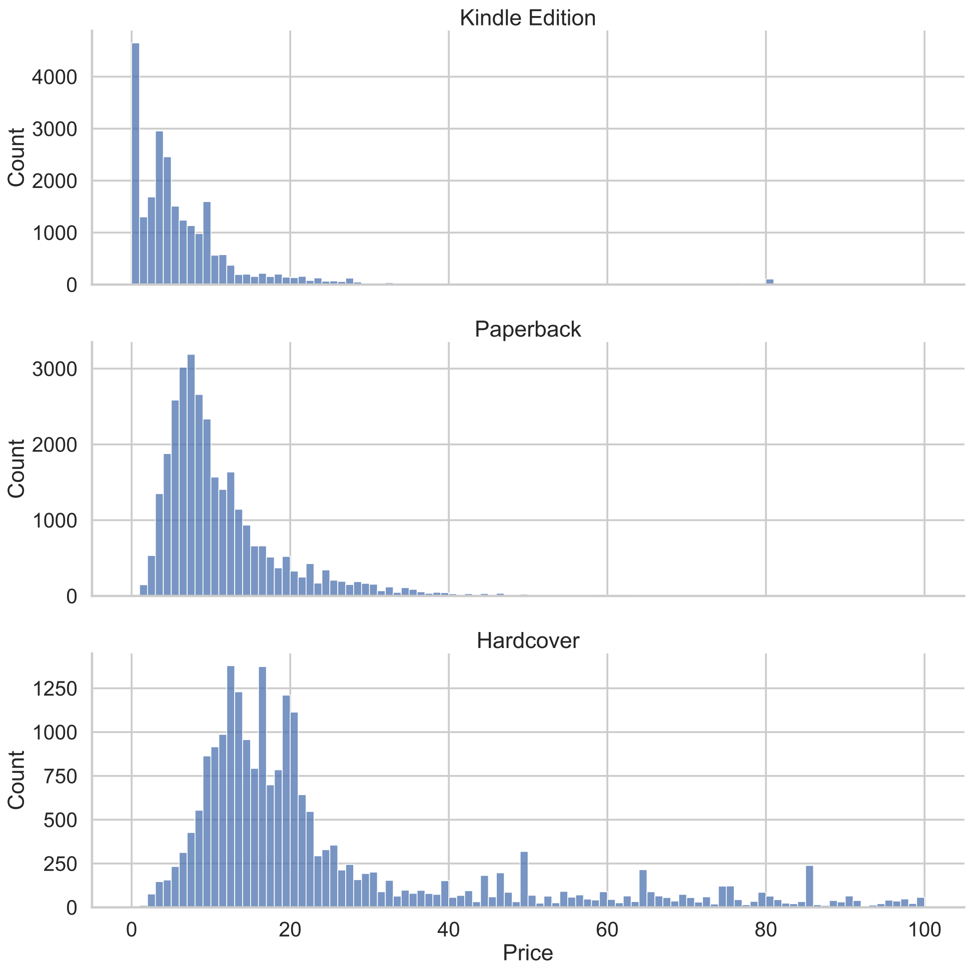

Figure 1: Distribution of retail prices by book format.

Figure 1 represents the frequency distribution of the retail prices for e-books (top), paperback (centre) and hardcover books (bottom) below £100. It is obvious that e-book prices are in a range between £0.25 and £10, paperback prices concentrate mostly in the £10-£20 interval and hardcover prices are even more expensive. While the distributions of e-books and paperbacks are more compressed, the hardcover book prices exhibit a higher volatility. Finally, all three price distributions have significant mass points at candidate focal points (e.g., £0.49 (e-books), £9.99 (paperback) and £15.99 (hardcover)).

Table 2 presents the descriptive statistics for Figure 1. With an average price of £8.34, e-books are the cheapest of the three book formats, followed by paperback (£12.34) and hardcover books (£28.54). The high standard deviation for hardcover books confirms its high volatility, which we have already detected in Figure 1. In overall, the descriptive statistics on book formats suggest that hardcover books exhibit the highest quality of the three formats and confirm the results from Li (2019), who has found that e-books and paperbacks are closer substitutes than e-books and hardcover books.

Format	count	mean	std	min	25%	50%	75%	max
Hardcover	22,647	28.54	42.17	1.05	12.99	18.54	28.96	1575.0
E-book	23,991	8.34	13.89	0.25	2.19	4.99	9.18	467.8
Paperback	30,870	12.35	18.86	0.31	6.55	9.01	13.95	1899.0

Table 2: Retail prices grouped by book format.

The e-books of the individual publishers are sold under different pricing arrangements. Amazon states on its product pages whether the respective publisher has set the price of an e-book. The Figures A.1 and A.2 in Appendix A present examples of this by showing screenshots for the e-book *Elon Musk: How the Billionaire CEO of SpaceX and Tesla is Shaping our Future* as well as for the e-book *Pulse*. In Figure A.1, it can be seen in the first box on the right hand side of the Amazon webpage that the 'price was set by the publisher' so that this is an example for the usage of the agency model. On the contrary, in Figure A.2 this information is missing, which means that Amazon sets the retail price for this e-book and it represents an example for the wholesale model.

Figure 2 visualizes the distribution of e-book retail prices by the different publishers. Prices are obviously more dispersed for book titles published by Bloomsbury and Oxford, whereas the other major publishers mostly have books in the range up to £20. The group of smaller publishers has a significantly larger fraction of e-books in the cheapest price range (about £0.49). In Appendix B, we will present this figure combined with assigned book genres.

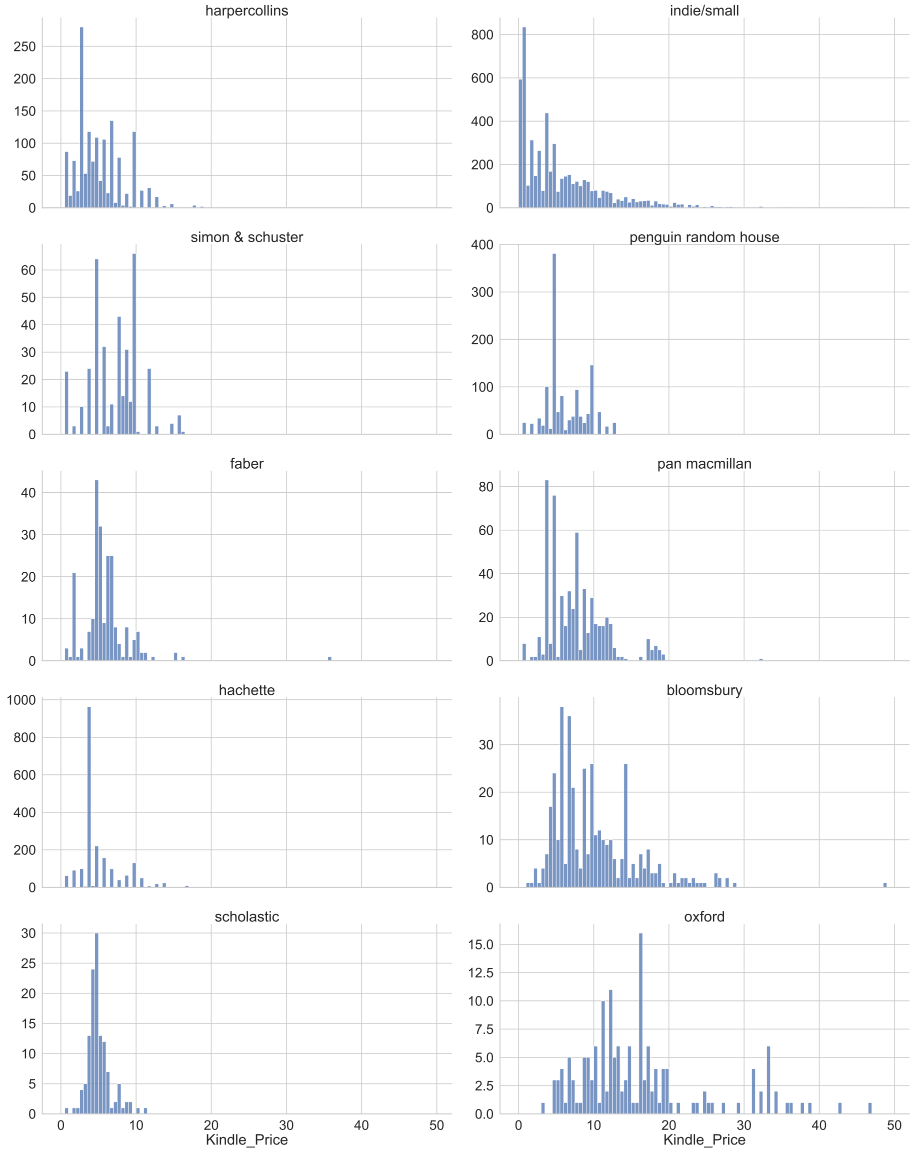

Figure 2: Prices for e-books grouped by publishers. The interval size for each bar is 1 Pound. For illustration purposes the figures are censored at 50 Pound.

In Table 3, we illustrate the distribution of the pricing arrangements used by the individual publishers for the e-books in our sample. If the number in the column 'Agency' takes the value one, the respective book titles are sold under the agency model, otherwise the titles are sold under the wholesale model. The table shows that all e-books in our dataset published by the large publishing houses Hachette, HarperCollins, Penguin Random House and Simon&Schuster are sold under the agency model on Amazon. Pan Macmillan has some American imprints which still use the wholesale model for their e-books. For the e-books in our sample published by Bloomsbury, Faber, Oxford, Scholastic and most of the smaller publishers, only the wholesale model is used.

Publisher	Agency	Amount	Percentage	Mean Price
Bloomsbury	0	397	100%	10.37
Faber	0	223	100%	5.84
Hachette	1	2,073	100%	5.31
Harper Collins	1	1,469	100%	5.42
Small Pub.	0	4,155	77.37%	6.81
	1	1,215	22.63%	2.32
Oxford	0	161	100%	16.71
Pan Macmillan	0	285	50.35%	9.75
	1	281	49.65%	5.81
Penguin Random House	1	1,240	100%	6.56
Scholastic	0	126	100%	5.01
Simon & Schuster	1	376	100%	7.45

Table 3: Distribution of the agency variable by publishers.

Finally, we want to illustrate the relationship between retail prices for e-books and their book sales rank on Amazon, which is illustrated in Figure 3 by using a scatter plot with a simple regression line. Obviously, there is a positive relation between the retail price and the rank of an e-book in our sample since the regression line has a positive slope. This finding in our sample is in line with the study of Fishwick (2008), who states that 'substantial discounts' (p. 370) have become prominent for bestselling books in the British book market after the abandonment of the Net Book Agreement in 1997.

With respect to the book sales rank, we have to stress how the rank on Amazon is determined. According to sources of Amazon, the ranks are internally updated hourly, but it does not appear immediately. The rank includes current and *all* past sales with higher weights on current sales.[11] This information on the definition of book ranks on Amazon will be important for our estimation approach because the price today might be affected by current sales or past sales, but not necessarily vice versa.

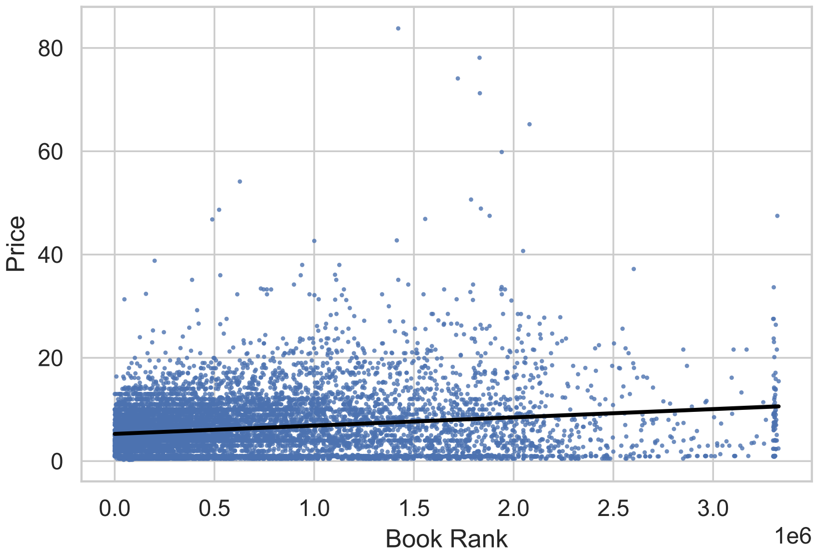

Figure 3: Relation of Amazon book ranks for e-books and retail prices.

[11]See https://kdp.amazon.com/en_US/help/topic/G201648140. (Last accessed: July 15, 2023)

4. Empirical Analysis

In this chapter, we present our empirical analysis. We first describe our estimation strategy in Section 4.1. In Section 4.2, we present the results of our OLS estimation, in which we estimate the effect of the pricing arrangement on the retail price of an e-book. The propensity score matching is outlined in Section 4.3.

4.1. Estimation Strategy

The goal of our study is to analyze the effect of the pricing arrangement on the retail prices of e-books sold on *Amazon.co.uk*. Therefore, we use publisher and book genre variation to estimate the price effect of e-books sold under the agency model. Before turning to the presentation of our estimations, we formalize the hypothesis that is to be tested. If there was no difference between the two types of vertical contracts (agency and wholesale) with respect to the retail price of an e-book, the ability for a publisher to set the final consumer price should (*ceteris paribus*) not have any impact on the final consumer prices. Hence, the hypothesis to be tested is:

Hypothesis 0 ($\boldsymbol{H_0}$): *The retail price of an e-book is independent of the used vertical contract.*

If there will be a positive correlation between the agency model and the price of e-books, H_0 can be falsified and e-books sold under the agency model are more expensive on average. Observing a negative correlation would also lead to a falsification of H_0, but then e-books sold under the agency model would be cheaper on average.

In our baseline estimation approach, we use the standard hedonic modeling approach in the spirit of Rosen (1974), which relies on observing differences in market prices to infer the value or implicit price of underlying characteristics. Thus, we estimate the following log-log OLS model with heteroscedasticity-consistent standard errors:

$$p_i = \alpha_0 + \alpha_1 A_i + \alpha_2 G_i + \alpha_3 R_i + \alpha_4 RRP_i + W\theta \ + \eta_i. \tag{1}$$

In equation (1), the dependent variable p_i is the logarithm of the retail price for an e-book i sold on *Amazon.co.uk*. The treatment variable A_i is a dummy variable which takes the value one if an e-book is sold under the agency model and zero otherwise. The variable G_i is a categorical variable for the book genres and R_i is a continuous variable for the e-book sales ranks on Amazon. The variable RRP_i reflects the recommended retail price of book title i. All other book-specific covariates are included in the matrix W (see Table 1 in Section 3.1).

However, there may be two endogeneity issues in regression equation (1), which might distort our estimated coefficients. First and foremost, our treatment variable A_i might be endogenous due to selection effects. Specifically, e-books sold under agency contracts might differ from e-books sold under wholesale contracts in ways that were directly related to their retail prices. Hence, it was probably not a random process whether an e-book is sold under the agency or the wholesale model. To eliminate this potential selection bias, we apply a matching procedure in Section 4.3 and perform two robustness checks in Section 5.

Second, the rank of an e-book R_i does not only affect the retail price of an e-book p_i but also does the retail price reflect the demand side and, therefore, affect the rank of an e-book which mirrors its sold quantity. As already explained, e-book ranks on *Amazon.co.uk* are internally determined by overall weighted sales ranks. Thus, ranks might be driven by the quantities sold today but the relation is ambivalent since the rank is also affected by quantities sold in the past. The e-book prices can be affected by current and past sales but the impact of prices on total (current and past) sales is not so clear. Nevertheless, we cannot clearly reject the endogeneity issue due to a potential reverse causality. To resolve this potential source of endogeneity between the e-book price and its sales rank, we will also present an instrumental variable approach in Appendix C.

4.2. OLS Estimation

The results of our log-log OLS regression model are outlined in Table 4. We estimate four different specifications in this estimation approach: in the first column of Table 4, we present a naive OLS regression model without genre effects, in column (2) we include the (LDA) genre fixed effects[12], in the third column we use the Amazon genre information as a control variable

[12] We describe the process to generate book genres by using an LDA approach in Appendix B. Thereby, we have identified a topic for every e-book title based on the largest probability assigned by the LDA.

and, finally, column (4) also includes an interaction term between the agency variable and the Amazon sales rank.

It is obvious that there is a negative and significant effect of the pricing arrangement *Agency* on the retail price of e-books in the specifications (1)–(3). According to the amount, the effect is between 18.3% and 18.9% depending on the exact specification. For the regression in column (1), an e-book which is sold under the agency model on *Amazon.co.uk* is approximately 18.5% cheaper than an e-book which is sold under the wholesale model on average.[13] Including the (LDA) genre fixed effects into the regression (see column (2) in Table 4) increases the agency effect to 18.9% (on amount), while the effect gets slightly lower when using the Amazon genre information (see column (3)). The inclusion of an interaction term between the agency variable and the book sales rank (see column (4)) leads to a significant positive coefficient of the variable *Agency*, while the interaction term has a negative sign. This result reveals that the negative mean price effect of the agency model is driven by e-books with higher sales ranks.

[13] To calculate the exact effect of the dummy variable *Agency* on the dependent price variable, the formula $100 \times (\mathrm{e}^{\beta} - 1)\%$ must be used.

	Dependent Variable: log Price			
	(1)	(2)	(3)	(4)
Agency	-0.205^{***}	-0.210^{***}	-0.202^{***}	0.845^{***}
	(-0.0001)	(0.011)	(0.011)	(0.086)
log sales rank	0.053^{***}	0.060^{***}	0.051^{***}	0.108^{***}
	(-0.0002)	(0.004)	(0.004)	(0.005)
log Kindle Size	0.045^{***}	0.050^{***}	0.027^{***}	0.051^{***}
	(-0.0001)	(0.004)	(0.004)	(0.004)
log star rating	0.531^{***}	0.491^{***}	0.436^{***}	0.494^{***}
	(0.0005)	(0.048)	(0.048)	(0.047)
No. expert reviews	0.015^{***}	0.015^{***}	0.018^{***}	0.009^{**}
	(-0.0001)	(0.004)	(0.004)	(0.004)
log RRP	1.091^{***}	1.076^{***}	1.080^{***}	1.057^{***}
	(-0.0001)	(0.011)	(0.012)	(0.011)
Date Retail	0.010^{***}	0.013^{***}	0.008^{***}	0.010^{***}
	(-0.00002)	(0.003)	(0.003)	(0.003)
Bestsellers	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.00000)	(0.0004)	(0.0004)	(0.0004)
WeekInChart	0.007^{***}	0.007^{***}	0.007^{***}	0.006^{***}
	(-0.00002)	(0.002)	(0.002)	(0.002)
Agency x log sales rank				-0.085^{***}
				(0.007)
Constant	-1.856^{***}	-1.970^{***}	-1.745^{***}	-2.522^{***}
	(0.004)	(0.075)	(0.074)	(0.082)
Genre	No	Yes	Amazon Genres	Yes
Robust F Statistic	1285.3022	751.67	627.4533	892.6006
Observations	12,001	12,001	12,001	12,001
Adjusted R^2	0.602	0.610	0.615	0.615

Robust Standard Errors in parentheses — *p<0.1; **p<0.05; ***p<0.01

Table 4: OLS estimation approach. Dependent variable is the logarithm of the retail price for e-books sold on Amazon.

The estimated coefficients for the other explanatory variables shown in the table are very similar for both specifications. The sign of the variable *log sales rank* indicates that e-books with higher sales ranks are sold at higher prices, which we have already demonstrated descriptively in Figure 3 of Section 3.2 and which confirms the results of Fishwick (2008), whereupon bestsellers are sold cheaper in the UK due to 'substantial discounts'. However, we will discuss potential endogeneity issues concerning this control variable in Appendix C. Moreover, there is a significant and positive relation between the RRP of an e-book and its retail price (see variable *log RRP*). We can further observe that the memory space of an e-book (given in KB) has a positive and significant price effect (see variable *log Kindle Size*).

The results presented in Table 4 also indicate that consumer and expert recommendations seem to drive the price of an e-book because the consumer star rating (*log star rating*) as well as the number of expert reviews (*No. expert reviews*) have a positive and significant effect in our regressions. Moreover, the time since the publication of a book title (in years) (variable *Date Retail*) seems to have a positive effect on e-book prices.

Lastly, two covariates remain in Table 4 which control for the author's quality. The explanatory variable *WeekInChart* reflects the average number of weeks former bestsellers of an author have last in the bestseller charts of the *Sunday Times* and the continuous variable *Bestsellers* exhibits the number of bestselling book titles an author has written historically. As expected, both variables have a positive and significant effect on the retail price of an e-book in all four specifications, which can be interpreted as quality signals increasing the price of a book title.[14]

However, the results of our OLS estimation in Table 4 might be biased and inconsistent due to endogeneity issues regarding our treatment variable *Agency* (see Section 4.1 for a discussion). Hence, we will apply a matching

[14]The variables *WeekInChart* and *Bestsellers* are based on a historical *Sunday Times Bestseller* list. The matching process was conducted via Python's Fuzzy Matching.

procedure in Section 4.3 to reduce this potential selection bias. Beyond, we will follow an instrumental variable approach in Appendix C to resolve the potential source of endogeneity of our important control variable book sales rank.

4.3. Propensity Score Matching

A simple OLS estimation would allow us to identify the effect of different pricing arrangements if the treatment was random. Yet, it was probably not a random process whether an e-book is sold under the agency or wholesale model so that there is a selection bias when using a standard OLS estimation approach. Hence in a first step, we seek to reduce this selection bias by relying on the propensity score matching (Rubin, 1977; Rosenbaum and Rubin, 1983). Thereby, we identify appropriate treated and control e-books through a matching procedure.

In particular, we identify those e-books that are not sold under the agency model but that had ex-ante the same probability of being sold under the agency model as those that are actually sold under the agency model. To implement the propensity score matching, we first run a logistic regression to recover the likelihood that an e-book is sold under the agency model based on its observable characteristics and use the predicted values from that estimation to collapse those covariates into a single scalar called the propensity score. Second, we match an e-book sold under the wholesale model which is as similar as possible to the considered e-book sold under the agency model based on this propensity score.

The propensity score is the selection probability conditional on the confounding variables ($p(X) = Pr(D = 1|X)$) and relies on two identifying assumptions. The first assumption is the conditional independence assumption, which requires that the outcome variable is independent from the treatment, conditional on the propensity score. This means we should only include variables that are expected to simultaneously influence both the treatment and the outcome. The second assumption is called common support assumption and simply means that for any probability there must be units in both the treatment group and the control group.

Then, the model we estimate through a logit regression as a first step of the matching procedure is:

$$A_i = \alpha + X_i\beta + \eta, \tag{2}$$

where A_i is the probability that an e-book i is sold under the agency model and X_i is a vector of e-book characteristics (see the explanatory variables in equation (1)).

	Agency
log RRP	-0.212***
	(-5.39)
log sales rank	-0.173***
	(-12.42)
log star rating	-0.255
	(-1.91)
No. expert reviews	0.375***
	(21.09)
log Kindle Size	0.128***
	(7.72)
WeekInChart	0.0113
	(1.28)
Date Retail	-0.0181
	(-1.80)
Bestsellers	0.0183***
	(3.46)
Constant	1.078***
	(3.77)
Observations	12,001
Pseudo R-squared	0.0924
Chi-square	1,524.7
Genre	Yes

t statistics in parentheses

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

Table 5: Logistic regression with *Agency* as the dependent variable.

In Table 5 we report the estimates for equation (1) on which the propensity scores are estimated. The frequency distributions of the propensity scores for the treated and untreated e-books are presented in Figure 4. As the Figure shows, the frequency distributions of the propensity scores for the two groups of e-books are very similar, indicating that there is a good set of e-books sold under the wholesale model that can be matched with e-books sold under the agency model based on their characteristics. Only a few e-book titles drop out of the sample after the matching process (indicated by the blue and yellow bars in Figure 4).

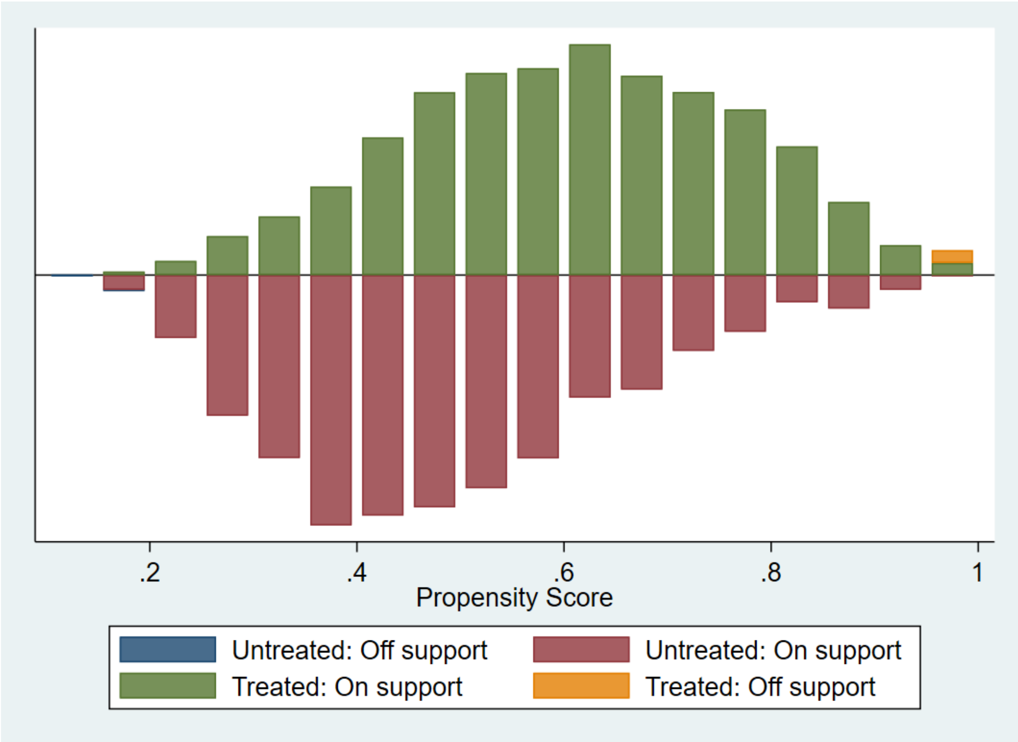

Figure 4: Propensity scores of a logistic model with Gaussian kernel fitting model. We have 11,955 observations which are on support. We present the covariate balance in Appendix Figure A.3.

In a second step, we use the estimated propensity scores to produce a sample where only the matched pairs of e-books remain to estimate the average treatment effect (ATE) of the agency model on the retail price of e-books. The empirical model that we employ is reported in equation (3), where p indicates two paired e-books (given by the matching procedure):

$$\Delta p_p = \alpha_0 + \alpha_1 A + \alpha_2 \Delta X_p + \epsilon_p. \tag{3}$$

Thus, Δp_p is the difference in the retail price between a treated e-book and a non-treated (control) one. The treatment variable A is equal to one if an e-book is sold under the agency model and zero otherwise. ΔX_p gives the differences in the e-book characteristics between the treatment and the control group.

	(1) Neighbor5	(2) Neighbor10	(3) Kernel	(4) Caliper
Agency	-0.227***	-0.233***	-0.235***	-0.240***
	(-14.94)	(-16.07)	(-20.28)	(-19.23)
OnSupport	11,955	11,955	11,955	11,955

t statistics in parentheses

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

Table 6: Propensity Score Matching. Standard Errors calculated by bootstrap with 101 random draws. The different models Neighbor5, Neighbor10, Kernel and Caliper refer to a 5-nearest neighbor, 10-nearest neighbor, a Gaussian kernel and a caliper matching, respectively. The caliper radius is determined by 0.2 times one standard deviation from the propensity scores derived from the logistic regression in Table 5 (see Austin, 2011, for optimal caliper length).

Table 6 reports the average treatment effect (ATE) from equation (3). We apply four different matching methods (two nearest neighbour covariate matching procedures, kernel and caliper) to deal with the issue of possible non-randomness in the treatment. The agency model still has a significant and negative effect on the retail price of e-books. The effect ranges between 20.3% and 21.3% depending on the used matching procedure.[15]

[15] Even though this seems to be a relatively large effect at first glance, it is put into perspective when looking at an example in absolute terms. The median retail price of an e-book in our dataset is £4.99 (see Table 2). Supposing that this median e-book is sold under the wholesale model (at £4.99), a negative agency effect of 20% in our estimation approach implies that the same e-books would be sold for £3.99 under the agency pricing model in a counterfactual scenario.

5. Robustness Checks

To check the robustness of our results presented in the previous section, we apply two further estimation approaches. In Section 5.1, we use data on the paperback format to apply an alternative OLS estimation approach. Following, we present a double machine learning (DML) approach in Section 5.2.

5.1. Format Estimation

The coefficient for the agency variable in our OLS estimation approach might be biased since we compare different book titles with each other. An ideal analysis would compare the same title offered to similar consumers under two separate vertical contracts. However, a certain e-book title is either sold under the agency or the wholesale model so that the counterfactual scenario is unknown. As an alternative, we run a robustness check in which we compare price differences between the digital and the paperback format for the same book title. Since print books are generally sold under wholesale contracts in the UK, this allows us to identify the price effect of the agency model on e-book prices (by also controlling for format differences).

To implement this robustness analysis, we only keep e-books sold under the agency model with an equivalent paperback version in our sample, which reduces our data set to 7,532 observations. Then, we estimate the following OLS model:

$$p_{i,j} = \alpha_0 + \alpha_1 A_{i,j} + \alpha_2 F_{i,j} + W\theta \; + \epsilon_{i,j}. \tag{4}$$

The dependent variable $p_{i,j}$ in equation (4) is the logarithm of the retail price for a book title i in format j. Again, the treatment variable $A_{i,j}$ is a dummy variable with value one if a book is sold under the agency model and zero when the wholesale contract is used. The format fixed effect $F_{i,j}$ is equal to one for the paperpack version and zero for the digital one. All other covariates are captured by the matrix W.

	Dependent Variable: log Price		
	(1)	(2)	(3)
Constant	-0.619^{***}	-0.396^{***}	-0.263^{***}
	(0.035)	(0.037)	(0.041)
Agency	-0.214^{***}	-0.272^{***}	-0.223^{***}
	(0.018)	(0.021)	(0.020)
Format	0.471^{***}	0.434^{***}	0.442^{***}
	(0.014)	(0.015)	(0.014)
log star rating	0.253^{***}	0.207^{***}	0.129^{***}
	(0.045)	(0.045)	(0.042)
No. expert reviews	0.040^{***}	0.032^{***}	0.029^{***}
	(0.005)	(0.006)	(0.006)
log RRP	0.932^{***}	0.930^{***}	0.841^{***}
	(0.012)	(0.012)	(0.015)
Date Retail	0.007^{**}	0.005	0.005
	(0.004)	(0.004)	(0.003)
Bestsellers	-0.0004	-0.001^{***}	-0.0004^{*}
	(0.0003)	(0.0002)	(0.0002)
R-squared	0.593	0.612	0.635
Adj. R-squared	0.593	0.611	0.633
Number of observations	7,532	7,532	7,532
Publisher	No	Yes	Yes
Genre	No	No	Yes

Robust Standard Errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table 7: Robustness check with paperback books (log-log OLS). Dependent variable is the logarithm of the retail price for digital and paperback versions of book titles sold on Amazon.

Results of equation (4) are reported in Table 7. Comparing these results to those reported in Table 6 we observe that the agency effect is very similar, even though it is larger in amount here (between 19.3 and 23.8%). The estimated coefficients for the variable *Format* imply that the paperback versions are significantly more expensive than the digital ones, confirming our descriptive statistics in Section 3.2 (see Figure 1 and Table 2). The coefficients of the remaining covariates in Table 7 are very similar to our main estimations in Section 4.

5.2. Double Machine Learning (DML)

There are further methods beyond the established approaches in the standard econometric analysis for causal inference. We have already applied standard econometric methods as the OLS estimation and the propensity score matching to deal with econometric issues. Recent advances in machine learning approaches also offer a larger toolbox for empirical analyses in economics (e.g., Athey and Imbens, 2019; Athey, 2018, for a broad overview).

Due to recent developments in the machine learning literature there are many approaches, e.g. the DML technique, that gives the possibility to deal with common econometric issues as confounding variables or variable selection by using cross-validation and non-parametric models. This technique also allows to make use of non-standard modelling for relations between variables, like the independent variables have a specific, often assumed linear or quadratic effect on the dependent variable. It permits to use any arbitrary machine learning technique relying on algorithms to find a fitting model for some chosen score functions like the mean squared error.

The reason for relying on further non-parametric/semi-parametric regressions is to circumvent the imposition of a specific model structure. The algorithm of the machine learning models will choose the best fitting model under some restrictions or parametrization. We will then use regularized linear regression techniques like the least absolute shrinkage and selection operator (Lasso) or regression trees/forests. These methods help to compare our standard econometric approaches with models that are able to ignore irrelevant variables or include non-linear effects (Athey and Imbens, 2019). Moreover, this methods can help in a similar way like propensity score matching to deal with the underlying selection issue (e.g., Lee et al., 2010; Westreich et al., 2010; Knaus, 2021).

Therefore, we apply DML techniques and compare the results to our previous estimations to provide further robustness checks. From a prediction's perspective, the estimations will be split into three different regression models, which is proposed by Chernozhukov et al. (2017, 2018). The models use the DML framework to deal with high dimensional variables, non-parametric functional forms, or unobserved confounding variables which have to be addressed by, e.g., an instrumental variable approach, based on the *DML-Conditional-Average-Treatment-Effect-Estimator*.[16]

The equation system we estimate has the following general form (we drop the index i for each book), which stems from the partial linear regression

[16]For the estimation implementation we follow the Python Module *econml* provided by Battocchi et al. (2019) (see also `https://econml.azurewebsites.net/spec/estimation/dml.html#overview-of-formal-methodology`).

model of Robinson (1988):

$$
\begin{aligned}
p &= \theta A + q(W) + \varepsilon \\
A &= f(W) + \eta \\
s.t.\ \mathbf{E}[\varepsilon|W] &= \mathbf{E}[\eta|W] = \mathbf{E}[\varepsilon \cdot \eta|W] = 0,
\end{aligned} \tag{5}
$$

where p denotes the price of an e-book depending on the agency dummy variable A and function $q(W)$ depending on the covariates W. The agency dummy variable A is explained by a function $f(W)$ of these covariates W (similar to the logit specification in equation (2)). The variables η and ε represent stochastic error terms. The conditional expectation functions then can be solved by some non-parametric regressions:

$$
\begin{aligned}
q(W) &= \mathbf{E}[p|W] \\
f(W) &= \mathbf{E}[A|W].
\end{aligned} \tag{6}
$$

In a next step, the residuals of the price, $\tilde{p}$, and the residuals of the agency dummy variable, $\tilde{A}$, are computed by subtracting the fitted values (given by the regression tasks in (6)) from the actual price p and the actual agency dummy variable A:

$$
\begin{aligned}
\tilde{p} &= p - q(W) \\
\tilde{A} &= A - f(W)
\end{aligned} \tag{7}
$$

Finally, we use these residuals for the prices $\tilde{p}$ and the agency dummy variable $\tilde{A}$ to estimate a linear treatment effect θ that is unbiased based on the assumptions of Chernozhukov et al. (2018):

$$\tilde{p} = \theta\tilde{A} + \epsilon. \tag{8}$$

In Table 8, the column *Model* represents the applied functional form. The first entry in this column refers to the functional form of computing and predicting $\tilde{p}$ and the second one for classifying $\tilde{A}$. Therefore, *Lin-Logit* relies on OLS and a logistic regression, *Lin-Lasso* relies on an OLS and a logistic regression including L_1 penalty (called Lasso), *Lasso-ElasticNet* uses a Lasso and a logistic regression with a combination of L_1 and L_2 penalties (*Elastic Net)*, *Lasso-RFC* refers to a Lasso regression and a random forest classifier, *RFR-ElasticNet* combines a random forest and an *Elastic Net*, *RFR-RFC* uses a random forest for both stages and XGBoost relates to Extreme Gradient Boost for both stages.[17]

The column *Score* in Table 8 displays the mean squared error of the final stage. In the final stage, a simple linear regression is used to get the conditional average treatment effect. The hyperparameters for each model are chosen from a reasonable set and then we use $3-5$ cross-fold validation within Python's Sklearn GridSearch. Besides, we also do another 5-fold splitting in each estimation. The presented results outline the best

[17]There are many more possible estimation techniques but these are sufficient to highlight the stability of our results.

estimation (lowest score) for each model class.

Model	Agency	Std. Error	p-value	Score	Perc_Change
Lin-Logit	-0.1957	0.0112	0.0000	0.3060	-17.7707
Lin-Lasso	-0.1956	0.0112	0.0000	0.3060	-17.7686
Lasso-ElasticNet	-0.1955	0.0112	0.0000	0.3061	-17.7578
Lasso-RFC	-0.2165	0.0126	0.0000	0.3064	-19.4696
RFR-ElasticNet	-0.1607	0.0094	0.0000	0.2134	-14.8437
RFR-RFC	-0.1800	0.0112	0.0000	0.2135	-16.4753
XGBoost	-0.1389	0.0120	0.0000	0.2313	-12.9706

Table 8: Double Machine Learning Approach. The dependent variable is the e-book retail price and the treatment variable is *Agency*. The price effect of the agency arrangement for the respective model is given in the column *Agency* and represents the ATE. Column *Score* refers to the mean squared error. Note: The data has been mean centered with unit variance as Lasso requires this normalization.

The point estimates of the individual regression models are given in the column *Agency* and the relative percentage changes in relation to the intercept are presented in the column *Perc Change* of Table 8.[18] In overall, the DML techniques confirm the results of our main estimations presented in Section 4 and prove the robustness of our regressions, even if we rely on more flexible methods. For instance, e-books sold under the agency model are 16.5% cheaper (on average) than digital books sold under the wholesale model when using the regression model *RFR-RFC*.

6. Conclusion

In this paper, we provide evidence that e-books sold under the agency model on *Amazon.co.uk* are significantly cheaper than e-books sold under the wholesale model on average. Our results are based on an unique dataset

[18] Again, one can calculate the exact percentage change by using the formula $100 \times (e^{Agency} - 1)\%$.

containing many characteristics of an e-book. To measure the relationship between the retail price of an e-book and the used pricing arrangement, we rely on classical econometric techniques as well as on newer methods as the DML approach. We find an robust and statistically significant effect that e-books sold under the agency model are approximately 20% cheaper than digital books sold under the wholesale model.

The results of our empirical analysis are in line with many theoretical papers studying the price effect of the agency model. Those theoretical analyses argue that retail prices for e-books sold under the agency model are lower due to the elimination of double marginalization (Lu, 2017), a lock-in effect exploited by retailers (Johnson, 2020), the monopolistic power of retailers over a complementary device as it is the case for Amazon when e-books could only be read on a Kindle device (Gaudin and White, 2014) or because agency selling is just more efficient than the wholesale model and leads to lower retail prices (Abhishek et al., 2016). Our results also match to the model-theoretical analysis from Foros et al. (2017) if one assumes that competition should be greater among publishers than among retailers, which most likely is the case due to the quasi-monopolistic power of Amazon.

To the best of our knowledge, this paper is the first empirical analysis estimating the price effect of the agency model for e-books by not only incorporating bestselling titles but also *long tail* book titles. Besides, in contrast to previous empirical analyses regarding the retail price of e-books, we apply an LDA approach to determine book genres. Nevertheless, a limitation of our approach is that we use cross-sectional data instead of panel data so that we cannot control for any dynamic effects on the retail price of e-books. Moreover, we only include one online platform, namely Amazon, instead of comparing various online retailers. Even though Amazon has a relatively high market share for e-books in the UK (see footnote 5), competition between the individual retailers very likely also has an effect on e-book prices.

The dynamic effect of the agency model on e-book prices including best-selling as well as *long tail* book titles remains an open question. Future research should concentrate on panel data to address such dynamic effects of different vertical contracts. Thereby, also other online platforms selling e-books should be included in such an analysis to identify effects between the online retailers. Finally, the long-run effect of the agency model on consumer welfare is an interesting research area. Consumer welfare does not only depend on the e-book prices, but also on other factors such as the number, variety, and quality of book titles written and published.

Declarations

Ethical Approval: Not applicable.

Competing interests: The authors have no competing interests to declare that are relevant to the content of this article.

Authors' contributions:

Maximilian Maurice Gail: Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Visualization, Writing – Review & Editing.

Phil-Adrian Klotz: Conceptualization, Methodology, Validation, Formal analysis, Writing – Original Draft, Supervision.

Funding: No funding was received to assist with the preparation of this manuscript.

Availability of data and materials: Data will be made available on request.

References

Abhishek, V., Jerath, K., Zhang, Z.J., 2016. Agency selling or reselling? channel structures in electronic retailing. Management Science 62, 2259–2280.

Aguiar, L., Waldfogel, J., 2018. Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music. Journal of Political Economy 126, 492–524.

Athey, S., 2018. The impact of machine learning on economics, in: The economics of artificial intelligence: An agenda. University of Chicago Press, pp. 507–547.

Athey, S., Imbens, G.W., 2019. Machine learning methods that economists should know about. Annual Review of Economics 11, 685–725.

Austin, P.C., 2011. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharmaceutical statistics 10, 150–161.

Battocchi, K., Dillon, E., Hei, M., Lewis, G., Oka, P., Oprescu, M., Syrgkanis, V., 2019. EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation. URL: https://github.com/microsoft/EconML. version 0.12.0.

Brynjolfsson, E., Hu, Y., Smith, M.D., 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. Management science 49, 1580–1596.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. The Econometrics Journal 21.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., et al., 2017. Double/debiased machine learning for treatment and causal parameters. Technical Report. Massachusetts Institute of Technology.

Condorelli, D., Galeotti, A., Skreta, V., 2018. Selling through referrals. Journal of economics & management strategy 27, 669–685.

Davies, S., Coles, H., Olczak, M., Pike, C., Wilson, C., 2004. Benefits from competition:

Some illustrative uk cases. DTI ECONOMICS PAPER NO. 9 .

Dearnley, J., Feather, J., 2002. The uk bookselling trade without resale price maintenance an overview of change 1995–2001. Publishing research quarterly 17, 16–31.

Fishwick, F., 2008. Book prices in the uk since the end of resale price maintenance. International journal of the economics of business 15, 359–377.

Foros, Ø., Kind, H.J., Shaffer, G., 2017. Apple's agency model and the role of most-favored-nation clauses. The RAND Journal of Economics 48, 673–703.

Gaudin, G., White, A., 2014. On the antitrust economics of the electronic books industry. Available at SSRN 2352495 .

Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as data. Journal of Economic Literature 57, 535–74.

Gilbert, R.J., 2015. E-books: A tale of digital disruption. Journal of Economic Perspectives 29, 165–184.

Ippolito, P.M., 1991. Resale price maintenance: Empirical evidence from litigation. The journal of law and economics 34, 263–294.

Johnson, J.P., 2017. The agency model and mfn clauses. The Review of Economic Studies 84, 1151–1185.

Johnson, J.P., 2020. The agency and wholesale models in electronic content markets. International Journal of Industrial Organization 69, 102581.

Knaus, M.C., 2021. A double machine learning approach to estimate the effects of musical practice on student's skills. Journal of the Royal Statistical Society: Series A (Statistics in Society) 184, 282–300.

Larsen, V.H., Thorsrud, L.A., 2019. The value of news for economic developments. Journal of Econometrics 210, 203 – 218. URL: http://www.sciencedirect.com/science/article/pii/S0304407618302148, doi:https://doi.org/10.1016/j.jeconom.2018.11.013.

Lee, B.K., Lessler, J., Stuart, E.A., 2010. Improving propensity score weighting using machine learning. Statistics in medicine 29, 337–346.

Lenz, D., Winker, P., 2020. Measuring the diffusion of innovations with paragraph vector topic models. PLOS ONE 15, 1–18. URL: https://doi.org/10.1371/journal.pone.0226685, doi:10.1371/journal.pone.0226685.

Li, H., 2019. Intertemporal price discrimination with complementary products: E-books and e-readers. Management Science 65, 2665–2694.

Lu, L., 2017. A comparison of the wholesale model and the agency model in differentiated markets. Review of Industrial Organization 51, 151–172.

MacKay, A., Smith, D.A., 2017. Challenges for empirical research on rpm. Review of Industrial Organization 50, 209–220.

Mathewson, G.F., Winter, R.A., 1984. An economic theory of vertical restraints. The RAND Journal of Economics , 27–38.

McCallum, A.K., 2002. Mallet: A machine learning for language toolkit. http://mallet. cs. umass. edu .

Nelson, P., 1970. Information and consumer behavior. Journal of political economy 78, 311–329.

Perry, M.K., Porter, R.H., 1986. Resale Price Maintenance and Exclusive Territories in the Presence [of] Retail Service Externalities. Department of Economics, State University of New York at Stony Brook.

Poort, J., van Eijk, N., 2017. Digital fixation: the law and economics of a fixed e-book price. International Journal of Cultural Policy 23, 464–481.

Řehůřek, R., Sojka, P., 2010. Software Framework for Topic Modelling with Large Corpora, in: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, ELRA, Valletta, Malta. pp. 45–50.

Reimers, I., Waldfogel, J., 2021. Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. American Economic Review 111, 1944–71.

Rey, P., Stiglitz, J., 1988. Vertical restraints and producers' competition. Technical Report. National Bureau of Economic Research.

Rey, P., Stiglitz, J., 1994. The role of exclusive territories in producers' competition. Technical Report. National Bureau of Economic Research.

Robinson, P.M., 1988. Root-n-consistent semiparametric regression. Econometrica: Journal of the Econometric Society , 931–954.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy 82, 34–55.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.

Rubin, D.B., 1977. Assignment to treatment group on the basis of a covariate. Journal of educational Statistics 2, 1–26.

De los Santos, B., O'Brien, D.P., Wildenbeest, M.R., 2018. Agency pricing and bargaining: Evidence from the e-book market. Kelley School of Business Research Paper , 1–51.

De los Santos, B., Wildenbeest, M.R., 2017. E-book pricing and vertical restraints. Quantitative Marketing and Economics 15, 85–122.

Spengler, J.J., 1950. Vertical integration and antitrust policy. Journal of political economy 58, 347–352.

Steyvers, M., Griffiths, T., 2007. Probabilistic topic models. Handbook of latent semantic analysis 427, 424–440.

Thorsrud, L.A., 2020. Words are the new numbers: A newsy coincident index of the business cycle. Journal of Business & Economic Statistics 38, 393–409.

Tirole, J., 1988. The theory of industrial organization. MIT press.

Varian, H.R., 2014. Big data: New tricks for econometrics. Journal of Economic Perspectives 28, 3–28.

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, 31.

Westreich, D., Lessler, J., Funk, M.J., 2010. Propensity score estimation: neural networks, support vector machines, decision trees (cart), and meta-classifiers as alternatives to logistic regression. Journal of clinical epidemiology 63, 826–833.

Winter, R.A., 1993. Vertical control and price versus nonprice competition. The Quarterly Journal of Economics 108, 61–76.

Appendix A. Additional figures and tables

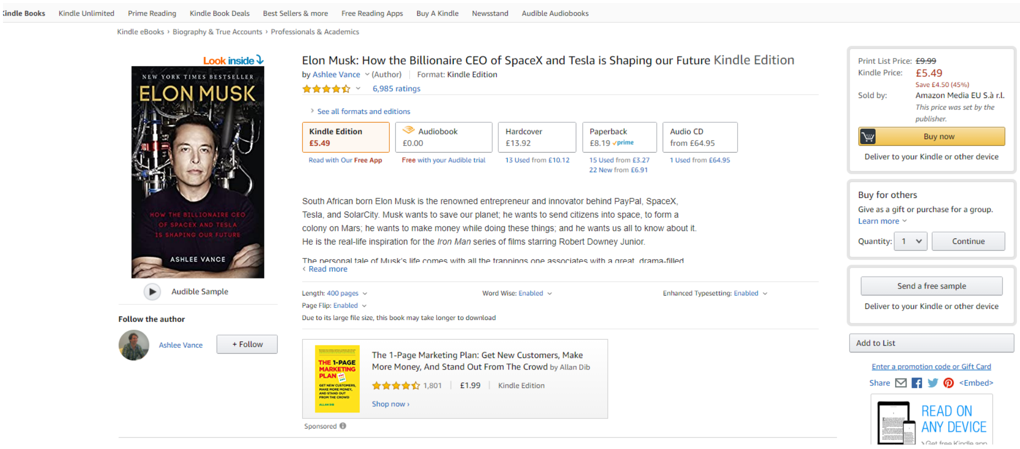

Figure A.1: Screenshot of *Elon Musk: How the Billionaire CEO of SpaceX and Tesla is Shaping our Future* (*Amazon.co.uk*).

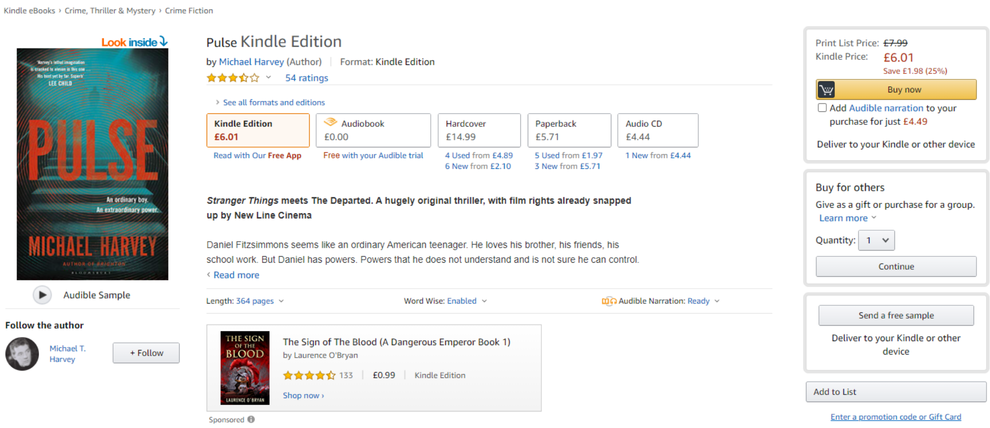

Figure A.2: Screenshot of *Pulse* (*Amazon.co.uk*).

	Publisher	Bloomsbury	Faber	Hachette	HarperCollins	Small Pub.	Oxford	Pan Macmillan	Penguin Random House	Scholastic	Simon & Schuster
Retail Price	mean	10.37	5.84	5.31	5.42	5.80	16.71	7.79	6.56	5.01	7.45
	std	7.02	3.21	2.86	3.20	6.07	10.12	3.86	2.70	1.58	3.27
Sales Rank	mean	554,409.87	326,982.37	262,402.04	427,905.75	778,429.39	937,724.43	497,575.36	365,604.06	597,692.75	555,114.93
	std	594,258.21	432,500.62	383,524.36	514,257.98	728,316.54	653,612.74	551,601.82	533,573.80	655,036.93	649,582.93
Star Rating	mean	0.90	0.85	0.89	0.89	0.86	0.90	0.89	0.88	0.93	0.90
	std	0.10	0.13	0.08	0.09	0.12	0.10	0.10	0.08	0.07	0.08
No. Customer Reviews	mean	51.10	60.80	119.85	106.88	80.50	17.19	124.70	134.59	76.33	142.89
	std	110.04	113.83	178.01	168.94	161.48	42.95	210.71	195.05	138.88	210.61
Pages	mean	292.57	280.60	388.78	342.59	286.86	412.85	336.34	319.01	231.84	367.91
	std	124.35	191.93	1949.17	330.58	212.83	207.28	121.63	140.63	113.11	256.70
Kindle Size	mean	13,078.55	2,849.00	14,134.92	8,288.26	8,626.52	9,995.83	12,118.73	20,466.37	35,684.97	15,118.97
	std	27,346.77	10,529.46	46,813.98	28,600.24	30,096.99	16,416.27	37,011.08	48,498.42	42,097.93	27,885.95
RRP	mean	15.50	10.28	13.07	12.04	12.08	35.29	13.61	14.59	8.31	13.85
	std	9.76	4.88	5.32	5.90	10.22	30.15	5.68	5.91	3.15	5.29
Date Retail	mean	1.64	2.41	1.81	2.09	1.93	2.97	1.66	2.18	2.11	2.13
	std	1.76	2.62	2.06	2.36	1.91	1.81	1.78	2.25	2.11	2.29
No. Expert Reviews	mean	1.74	2.43	2.33	1.25	1.34	1.14	2.24	1.71	0.65	1.35
	std	0.96	1.24	1.06	1.12	1.51	0.77	1.57	1.40	0.85	1.04

Table A.1: Summary Statistics

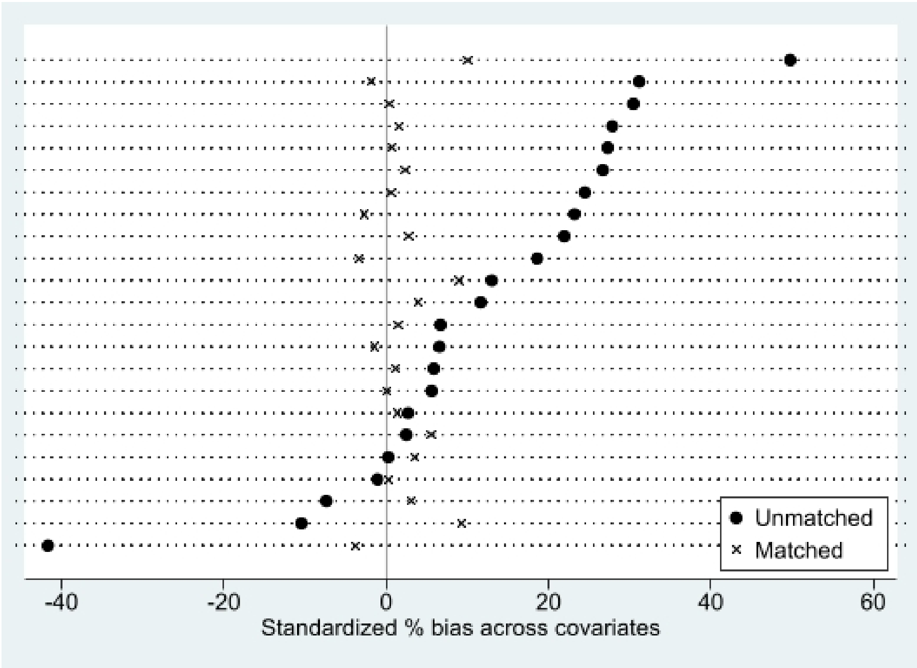

Figure A.3: Covariate Balance. Matched sample exhibits a mean (median) bias of 3.1 (2.3). Rubin's B and Rubin's R are 20.1 and 1.5, respectively.

Appendix B. Latent Dirichlet Allocation (LDA)

In the recent past, new technologies have made it possible to use text as data and, therefore, as an input to economic research. Text data, which is inherently high-dimensional, can capture relevant economic concepts not covered by "hard" economic data. In the last years, there has been an explosion of empirical economics research using text as data (e.g., see Larsen and Thorsrud (2019) for an Latent Dirichlet Allocation (LDA) approach or Lenz and Winker (2020) for paragraph vector topic modelling). We have decided to use an LDA approach to generate book genres and to assign every single book title from our dataset into one of these genres. Such a text mining approach is necessary because on the Amazon webpage the genre information is ambiguous and even not available for some book titles. For this purpose, we use the descriptions and expert reviews from the individual books in our dataset as text data input. We further rely on natural language processing (NLP) to extract the relevant information.

We apply several Python-Modules to clean and prepare the raw dataset.[19] Thereby, we remove common words and surnames, eliminate stop words, remove punctuation and pronouns as well as reduce all words to their respective word stems. We note here that around 45,819 unique tokens are kept after this filtering process.

This cleaned descriptions corpus is decomposed into book genres using

[19]Base module is gensim by Řehůřek and Sojka (2010) with a wrapper called Mallet, which is a Java-based open-source NLP text analytics tool (see McCallum (2002) or `http://mallet.cs.umass.edu/`).

the already mentioned LDA model. The LDA provides a statistical framework for the generation of documents based on topics. It is an unsupervised topic model that clusters words into topics/ genres, which are distributions over words, while at the same time classifying descriptions as mixtures of topics/ genres. The term "latent" is used because the words are intended to communicate a latent structure, namely, the subject matter (topic) of the description. The term "Dirichlet" is used because the topic mixture is drawn from a conjugate Dirichlet prior (Thorsrud, 2020).

The structure of the LDA model is as follows: the whole corpus is represented by M distinct documents (descriptions) and $N = \sum_{m=1}^{M} N_m$ is the total number of words in all documents. Assuming K latent topics/ genres, each topic is given by a probability vector $\boldsymbol{\phi_k} = (\phi_{k,1}, ..., \phi_{k,N})$ with $\sum_{n=1}^{N} \phi_{k,n} = 1$ indicating the probability that each word shows up in this topic. Further, each document $m \in \{1, ..., M\}$ contains all topics with different probabilities (weights) $\boldsymbol{\theta_m} = (\theta_{m,1}, ..., \theta_{m,K})$ with $\sum_{k=1}^{K} \theta_{m,k} = 1$. Both $\boldsymbol{\phi_k}$ and $\boldsymbol{\theta_m}$ are assumed to have conjugate Dirichlet distributions with hyper parameters (vectors) α and β, respectively.

Given $\boldsymbol{\phi_k}$ and $\boldsymbol{\theta_m}$, a document is generated by drawing for each word a topic $k \in \{1, ..., K\}$ according to the probabilities $\boldsymbol{\theta_m}$ and one word from the selected topic according to its distribution $\boldsymbol{\phi_k}$. This procedure is repeated until the length of the document is reached. To solve the LDA model, we a priori set $\alpha = 50$ and $\beta = 0.01$. The hyper parameter optimization is executed by using Gibbs simulations. Gibbs sampling (also known as alternating conditional sampling) is a specific form of Markov chain Monte Carlo and simulates a high-dimensional distribution by sampling on lower-dimensional subsets of variables where each subset is conditioned on the value of all others (e.g., Steyvers and Griffiths, 2007).

The sampling is done sequentially and proceeds until the sampled values approximate the target distribution. We set the number of sampling iterations equal to 1,000. Then, based on the coherence value across the estimated LDA models using smaller numbers of genres, we find that 12 topics/genres provide the best statistical decomposition of our book description corpus.[20] A detailed list of all 12 genres is presented in Table B.2.

Topic	Genre
0	History
1	Guidebook
2	Childen and Youth
3	Society Novel
4	Lifestyle
5	Crime Novels/Thriller
6	Politics
7	Historic Novel
8	Drama
9	Family Novel
10	Biography
11	Textbook

Table B.2: 12 different genres identified by our LDA approach.

One caveat of the LDA estimation procedure is that it does not give the

[20]For 12 different topics, the coherence value exhibits a local peak. We have also considered 9 and 17 different topics, but in the end there was no real change in the effects on the other variables.

topics/genres any names or labels. Thus, labels are subjectively given to each genre based on the most important words associated with each topic. In the most cases, it is conceptually simple to classify the genres. Besides, the exact labeling plays no material role in our empirical approach, it is just used as a convenient way of referring to the different topics instead of only using topic numbers.

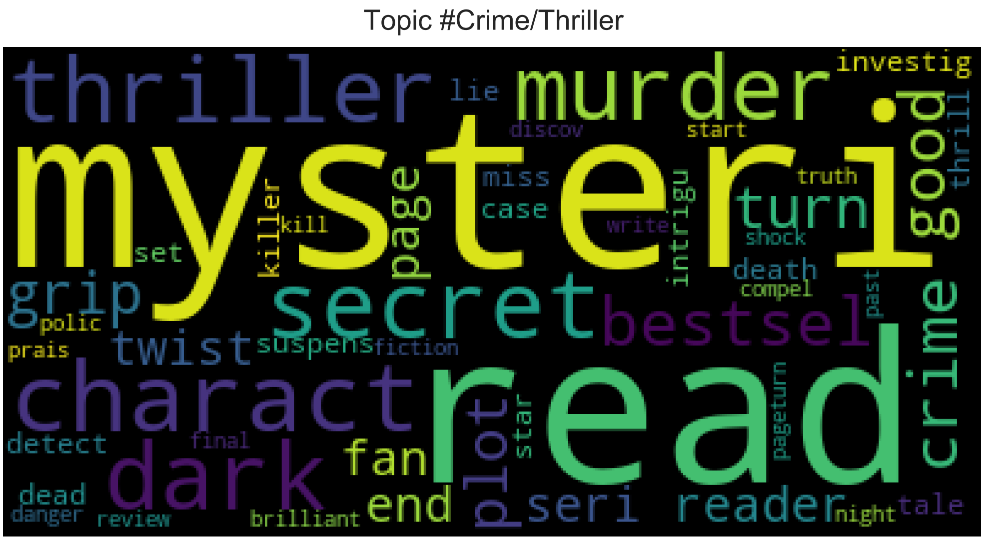

Figure B.4: Topic Crime Novel/Thriller from the LDA. Size is according to weight of word within the topic.

It is more important that the LDA decomposition gives a meaningful and easily interpretable genre classification of the book descriptions, which it does because our LDA approach identifies all important book genres and clearly delineates the topics. This is shown by the Figures B.4 and B.5, which are two examples of our 12 word clouds to visualize the genre distribution of words by assigned probabilities through the LDA. We have labelled the topic in Figure B.4 *Crime Novel/Thriller* and the genre in Figure B.5 *Politics*. The larger the size of a word in these clouds is, the higher is its weight within the respective topic.

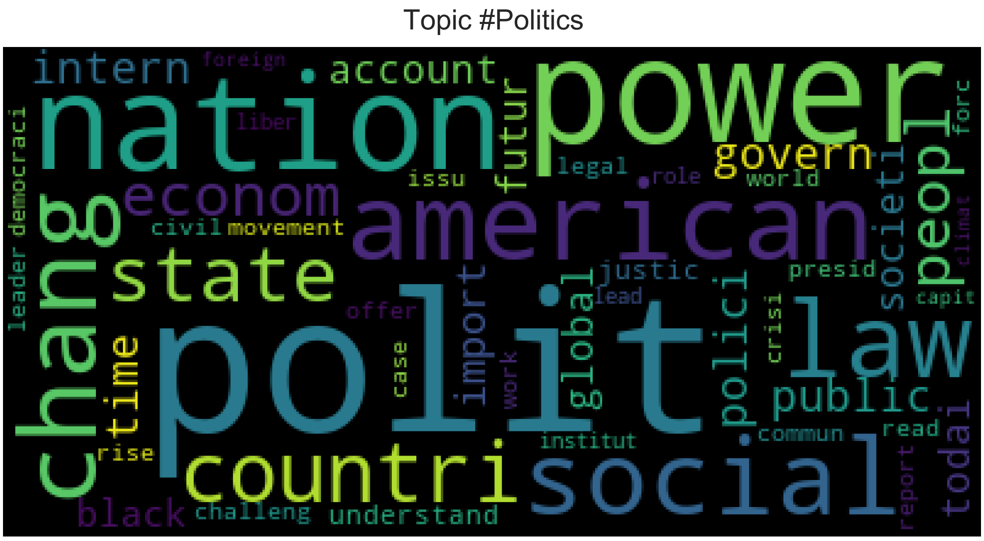

Figure B.5: Topic Politics from the LDA. Size is according to weight of word within the topic.

The LDA provides the ability to give each document (book description) from our corpus (data) a probability of being a respective topic. For the visualization purpose in Figure B.6, which is based on Figure 2, the largest probability value is chosen to highlight the distribution of topics over e-book prices and publishers. This distribution of prices by genres exhibits the high comparability between the several publishers in our dataset because they are not specialised in certain topics, but all publishers sell book titles from different genres. This is an advantage for our empirical approach because otherwise we would get multicollinearity issues and from an economic point of view we would fail in the sense that we could not compare publishers at all, if there were only specific topics from specific publishers.

However, it is obvious that the individual publishers have distinct main topics. For instance, Pan Macmillan primarily publishes fiction titles like crime novels, thrillers or society novels whereas HarperCollins has a focus on the genres family novel and drama. However, it is important not to take these topics at face value since the LDA assigns a probability to each individual topic.

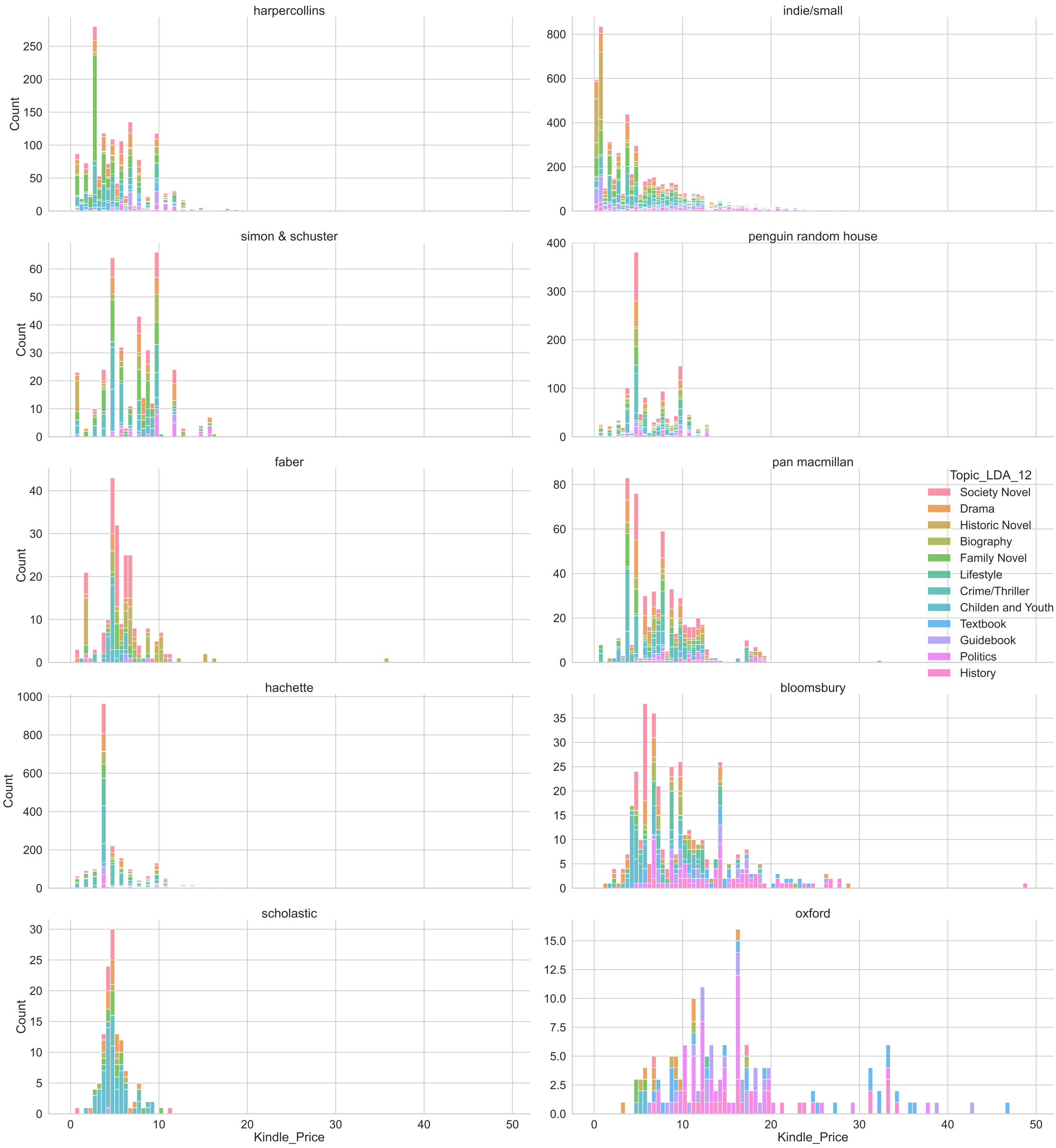

Figure B.6: Prices for e-books grouped by publisher and genre. The ordinate is scaled differently for each subplot.

Appendix C. IV Estimation

The results of our OLS estimation in Section 4.2 might be biased and inconsistent since the important control variable book sales rank might be endogenous. Hence, we will follow an instrumental variable approach in this section to resolve this potential source of endogeneity. We use the logarithmized number of customer reviews (*log no. customer reviews*) as an instrument for the book sales rank to avoid inconsistent estimates due to reverse causality.

Our instrumental variable *log no. customer reviews* is highly correlated with our endogenous regressor book sales rank (relevance condition) but should have no partial effect on the price of an e-book (orthogonality assumption). Customer reviews can enhance the awareness and information quality for a consumer and, thus, change the tendency for a consumer to purchase a book. However, the absolute number of customer reviews does not affect the purchasing decision of a consumer for a book title, but only surprisingly positive (negative) reviews can increase (decrease) the consumption of a given good (Reimers and Waldfogel, 2021). Hence, the absolute number of customer reviews should also have no partial effect on e-book prices, even though our instrument is highly correlated with the book sales rank (as it is an indicator for past sales).

Following the approach explained above, the linear projection in the first stage regression of our 2SLS estimation can be formalized as follows:

$$R_i = \beta_0 + \beta_1 A_i + \beta_2 P_i + \beta_3 G_i + \beta_4 P_i \times G_i + \beta_5 RRP_i + \beta_6 CR_i + W\theta + \xi_i. \quad \text{(C.1)}$$

In equation (C.1), the dependent variable R_i refers to the sales rank of book title i on *Amazon.co.uk*. We have already introduced most of the covariates used here in the context of our baseline estimation in equation (1). Our instrumental variable *log no. customer reviews* is displayed by CR_i.

The structural equation of our basic model then takes the following form:

$$p_i = \gamma_0 + \gamma_1 A_i + \gamma_2 P_i + \gamma_3 G_i + \gamma_4 P_i \times G_i + \gamma_5 RRP_i + \gamma_6 \hat{R}_i + W\theta + \varepsilon_i, \quad \text{(C.2)}$$

where the dependent variable p_i is the retail price of e-book i and the fitted values from the first-stage are captured by $\hat{R}_i$.

The regression results based on equation (C.2) are presented in columns (2) and (3) of Table C.3. In column (1), we again show the results of a (naive) OLS estimation when we control for the book genres. The two IV approaches in Table C.3 differ in terms of the added genre control variable (LDA generated or Amazon book genres). The results of our IV approach confirm that e-books sold under the agency model on *Amazon.co.uk* are on average significantly cheaper than book titles sold under the wholesale model. Compared to the OLS estimation result in column (1), the estimated coefficients for the variable *Agency* only differ in their magnitude as we find a negative effect of agency pricing between 17.4 and 17.9% now. Hence, we slightly overestimate the absolute effect of agency pricing when we ignore the endogeneity issue of the variable book sales rank. Also the effects of the other explanatory variables barely differ between the OLS and the IV estimation approaches, even though the effect of the book sales rank has become larger in the IV regressions.

	Dependent Variable: log Price		
	(1)	(2)	(3)
log sales rank	0.060***	0.093***	0.076***
	(−0.0002)	(−0.001)	(−0.001)
Agency	−0.210***	−0.197***	−0.191***
	(−0.0001)	(−0.0003)	(−0.0004)
log Kindle Size	0.050***	0.049***	0.026***
	(−0.0001)	(−0.0001)	(−0.0001)
log star rating	0.491***	0.514***	0.456***
	(0.001)	(0.0002)	(0.0001)
No. expert reviews	0.015***	0.021***	0.023***
	(−0.0001)	(−0.0002)	(−0.0002)
log RRP	1.076***	1.078***	1.082***
	(−0.0002)	(−0.0002)	(−0.0002)
Date Retail	0.013***	0.012***	0.008***
	(−0.00002)	(−0.00001)	(−0.00001)
Bestsellers	0.001***	0.002***	0.002***
	(0.00000)	(−0.00000)	(−0.00000)
WeekInChart	0.007***	0.010***	0.009***
	(−0.00002)	(−0.0001)	(−0.0001)
Constant	−1.970***	−2.413***	−2.076***
	(0.006)	(0.014)	(0.012)
Genre	Yes	Yes	Amazon Genres
Instrument	No	Reviews	Reviews
First Stage F-Statistic	-	3641.0689	3915.4707
Wu-Hausman Statistic	-	26.3868	17.0578
Observations	12,001	12,001	12,001
Adjusted R^2	0.610	0.607	0.614
Robust Standard Errors in parentheses			*p<0.1; **p<0.05; ***p<0.01

Table C.3: IV estimation results.

4 Pass-through of temporary fuel tax reductions: Evidence from Europe

Authors: Chiara Patricia Drolsbach, Maximilian Maurice Gail and Phil-Adrian Klotz

Published in: *Energy Policy 183 (2023): 113833.*

Energy Policy 183 (2023) 113833

Contents lists available at ScienceDirect

Energy Policy

journal homepage: www.elsevier.com/locate/enpol

Pass-through of temporary fuel tax reductions: Evidence from Europe

Chiara Patricia Drolsbach [a], Maximilian Maurice Gail [b], Phil-Adrian Klotz [b,*]

[a] Chair for Data Science and Digitization, Licher Strasse 74, Giessen, 35394, Germany
[b] Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Licher Strasse 62, Giessen, 35394, Germany

ARTICLE INFO

JEL classification:
H23
L13
L91
Q48

Keywords:
Fuel prices
Pass-through
Environmental taxes
Staggered diD

ABSTRACT

Several European countries have implemented temporary fuel tax reductions in 2022 to relieve the financial burden on their citizens. This paper is the first to provide estimates of the pass-through rates as well as the effect on retail margins for France and Italy. Further, it contributes to the recent literature on the fuel tax reduction in Germany. Using a unique data set containing daily consumer prices at service station chain level for gasoline and diesel, we employ a staggered Difference-in-Differences (DiD) design. Our main results imply that in the aggregate there was a full-shifting of the fuel tax reductions in all three countries. Nevertheless, in an event study design we find that the pass-through rates over time are heterogeneous between the countries and types of fuel. Depending on time, heterogeneous effects imply a full-shifting up to a minor over-shifting of the pass-through rates. These findings also have important implications for the effective design of unconventional fiscal policy as well as for competition policy in the fuel market.

1. Introduction

More than two years after the beginning of the COVID-19 pandemic many countries worldwide exhibited very high inflation rates. The reasons are a recovering demand in combination with ongoing supply chain problems as well as the war of aggression in the Ukraine. In April 2022, the inflation rate of Germany reached 7.4%, the highest rate since 1981. Other western countries have similar rates: the inflation rate of the whole EU has been 8.1%, the US even had a rate of 8.3% in April 2022. In autumn 2022, inflation in some countries has already risen to around 10%.[1]

In this situation several governments tried to relieve their citizens with tax reductions or transfer payments. On April 27, 2022, the German government announced a (second) stimulus package worth 14–16 billion Euro.[2] Beside new transfer payments and a cheap, nationwide public transport ticket (*€9 ticket*), it also included a temporary reduction of the energy tax rate from June 1 to August 31, 2022 at an estimated cost of 3.15 billion Euro.[3] Since the energy tax is levied on fuel products in Germany, this might also have had an effect on retail fuel prices. However, consumers only benefit from this regulation if the petroleum companies pass-through the tax reduction sufficiently.

Also other countries in the EU, such as France or Italy have implemented temporary measures in 2022. In France, the government introduced a fixed fuel discount between April 1 and August 31, which later has been extended until the end of the year 2022. The Italian government had already approved a subsidy program in March including a fuel tax reduction from March 22 until the end of April. Also this intervention has later been extended until December 31, 2022. Those government actions provide us with ideal exogenous shocks, which we can use as a natural experiment.[4]

In this paper, we estimate the pass-through rate and the effect on the retail margins of the temporary fuel tax reductions in the three largest countries of continental Europe (measured by GDP), namely France, Germany and Italy. Austria, Estonia, Lithuania, and Latvia are being selected as appropriate control countries for the purpose of this analysis because these nations did not introduce any comparable measures during the year 2022. Using a unique panel data set containing daily consumer prices for gasoline and diesel on service station chain level, we compute the pass-through rates and changes in the margins by employing a staggered Difference-in-Differences (DiD) approach.

Our results imply a heterogeneous passing over time of the fuel tax reductions depending on the country as well as on the type of fuel.

* Corresponding author.
E-mail address: phil.a.klotz@wirtschaft.uni-giessen.de (P.-A. Klotz).

1 See https://www.global-rates.com/de/wirtschaftsstatistiken/inflation/inflation.aspx. (Last accessed: October 19, 2022).

2 See https://www.bundesfinanzministerium.de/Content/DE/Pressemitteilungen/Finanzpolitik/2022/04/2022-04-27-zweites-entlastungspaket.html. (Last accessed: October 19, 2022). The package was approved by the German parliament on May 13.

3 See https://www.bundestag.de/dokumente/textarchiv/2022/kw20-de-energiesteuersenkungsgesetz-894664. (Last accessed: October 19, 2022).

4 Due to the circumstance that the interventions were introduced to ease the burden to consumers and with the intention to lower consumer prices, potential endogeneity issues will be discussed at the end of Section 5.

https://doi.org/10.1016/j.enpol.2023.113833
Received 3 November 2022; Received in revised form 8 September 2023; Accepted 21 September 2023
Available online 5 October 2023
0301-4215/

However, we find the following two key results. First, the average pass-through rates are very high so that there is a full-shifting of the temporary tax reductions, indicating highly competitive markets. Second, the estimated pass-through rates are on average higher for gasoline than for diesel,[5] which may result from the special situation on the European energy markets in 2022 following the Russian invasion of Ukraine and a relating high demand of (heating) diesel.

The results of our paper have implications for the effective design of unconventional fiscal policy and are also relevant for competition policy. Unconventional fiscal policy can only be effective in stimulating demand if consumers expect tax reductions to be passed through by firms. Besides, such fuel tax reductions also have distributional- and climate-economical effects. While the discount may act like a redistribution from bottom to top as particularly high-income consumers with large cars are benefiting, it is generally questionable whether subsidizing fossil fuels is a good idea in times of climate change.

The rest of the paper is structured as follows. Section 2 presents related literature, followed by a description of retail fuel markets in Section 3. We present our data set and descriptive statistics in Section 4. In Section 5, we explain our empirical strategy and then present the estimation results in Section 6. We conclude in Section 7 by discussing policy implications and limitations.

2. Related literature

Since gasoline markets are typically characterized by a very specific cyclical pricing pattern, academia as well as competition authorities are highly interested in analyzing this industry sector. The leading theory to explain price cycles in gasoline markets are Edgeworth price cycles. This theory has been formalized by Maskin and Tirole (1988) and assumes a dynamic oligopoly game where firms compete in prices and sell homogeneous goods. Starting at a supra-competitive price, firms undercut each other until the price reaches marginal costs. Given that there is no gain to lowering prices further, firms play a war of attrition. After one firm relents the price back to a high level, the other follow and the cycle begins anew (see Noel et al. (2011)).

In contrast to the literature mentioned above, other authors discuss the possibility of tacit collusion in gasoline markets. Since petrol stations can easily observe and monitor price changes as well as learn the price setting behavior of their competitors, an explicit agreement is not necessary to establish such an behavior. Evidence for collusion in gasoline markets has been found for Australia (Byrne and De Roos, 2019) and Norway (Foros and Steen, 2013). With respect to Germany, Dewenter et al. (2017) show that the introduction of the 'Markttransparenzstelle für Kraftstoffe' (market transparency unit for fuels, MTS-K)[6] in 2013 has increased both gasoline and diesel prices. Clark et al. (2023) find that algorithmic pricing has a significant effect on competition in the German gasoline market.

Another strand of the literature analyzes the effects of changes in the crude oil price on refined petroleum products. Here, most of the papers are focused on the oil-gasoline relationship. It has been shown that downstream prices seem to respond to increases in upstream prices more rapidly than their responses to decreases in upstream prices, so that there is a potentially asymmetric pass-through of increasing and decreasing costs ('rockets and feathers') (e.g., Grasso and Manera (2007) and Noel (2009, 2015)). In this context, similar studies explore the causes for this asymmetric relation between crude oil and gasoline. They identify refinery utilization rates and inventories as a main driver of those asymmetries (e.g., Kaufmann and Laskowski (2005) and Perdiguero-Garcia (2013))

Recent papers also analyze the pass-through of taxes and excise duties on fuel prices. In general, pass-through rates depend on consumer behavior as well as on competition parameters (e.g., Montag et al. (2021), Genakos and Pagliero (2022) and Harju et al. (2022)). The effect of tax changes on market prices primarily depends on supply and demand elasticities (Edgeworth, 1897). In a perfectly competitive market, the pass-through rate increases in the elasticity of supply and decreases in the elasticity of demand. However, if competition is not perfectly competitive, the pattern of tax incidence becomes more complex and several degrees of tax shifting are possible: under-, full- and over-shifting to consumers (see Appendix A). Besides, not only the horizontal market structure but also vertical market power has to be considered (Fuest et al., 2020).

Some empirical results indicate that the coefficient associated with taxes on gasoline prices is not statistically different from one (or slightly less than one) (e.g., Marion and Muehlegger (2011), Bello and Contín-Pilart (2012) and Li et al. (2014)). In contrast, other studies find that a higher percentage of a tax increase is passed to consumers than a tax reduction (Doyle Jr. and Samphantharak, 2008; Silvia and Taylor, 2014) or identify state-specific rates of pass-through (Kaufmann, 2019). Regarding the fuel tax reduction in Germany in 2022, results range between a partial pass-through to a full-pass-through (e.g., Dovern et al., 2023; Schmerer and Hansen, 2023; Kahl, 2023; Bernhardt et al., 2023; Fuest et al., 2022; Seiler and Stöckmann, 2023). Here we explicitly contribute for the case of Germany but also other not yet examined countries (France and Italy) utilizing a staggered difference-in-differences design.

3. The retail fuel market

The fuel market is characterized by a vertical structure, with refineries producing fuels from crude oil in the upstream market and selling them to fuel stations, which in turn distribute the fuels to end customers (downstream market). In our study, we focus on the analysis of retail prices on the service station chain level, however, an understanding of the upstream sector is still relevant, especially for the calculation of margins. During fuel production a barrel (42 gallons) of crude oil can be refined into 19 gallons of gasoline, 12 gallons of diesel and 13 gallons of other products.[7] In addition to crude oil, refineries also add other oils and liquids to the finished products that are sold to the petrol stations.

After significant increases in the European retail fuel prices at the beginning of 2022, several countries adopted measures with the aim of relieving consumers. For our analysis, we focus on the three largest economies in continental Europe that have introduced reductions of excise duties on fuel or similar measures, explicitly Germany, France, and Italy. To choose appropriate control countries for our staggered DiD approach, we need to find countries of the European Union (EU) which have not implemented any regulations in the fuel market in 2022. Table 1 presents an overview of policies introduced in all member states of the EU. It is obvious, that there are numerous overlaps in timing (i.e., measures came into force on the same day), which prevent a comparison. Apart from that, there are several countries that have chosen VAT reductions or price caps as policy measures, which also reduces comparability (due to varying magnitude of actual discounts over time). The consideration of all countries shows that by these criteria the majority of all EU countries are not eligible as control countries for our analysis.[8] Yet, Austria, Estonia, Latvia, and Lithuania, as countries that have not introduced any regulations, are considered suitable for comparison.

[5] Note: Our hypothesis tests indicate that there is no statistically significant difference between them (see Section 6.1).

[6] The MTS-K is an independent unit of the German competition authority. All petrol stations in Germany are legally bound to inform the MTS-K about price changes in real time (see https://www.bundeskartellamt.de/EN/Economicsectors/MineralOil/MTU-Fuels/mtufuels_node.html;jsessionid=0E947D4936B3B12872C630A4005CED95.2_cid378).

[7] See https://www.eia.gov/energyexplained/oil-and-petroleum-products/refining-crude-oil-inputs-and-outputs.php (Last accessed: October 19, 2022).

[8] Belgium, Croatia, Hungary, Poland, Portugal, and Slovenia have introduced regulations other than a fixed tax reduction/discount. Bulgaria, Czech

Table 1
Overview of fuel tax reductions in all EU member states. In the case of fuel tax reductions given values are excl. associated VAT reduction.
Sources: https://www.bruegel.org/dataset/national-policies-shield-consumers-rising-energy-prices (Last accessed: August 7, 2023).

Country	Country code	Type of measure	Date (mm/dd/yy)	Tax reduction		In sample?
				E5	Diesel	
Austria	AT	–	–	–	–	Control
Belgium	BE	VAT reduction + Fuel tax reduction	02/01 + 03/19/22	15% + 17.5ct/l	15% + 17.5ct/l	–
Bulgaria	BG	Fixed discount	07/09/22	13ct/l	13ct/l	–
Croatia	HR	Price cap	10/17/21			–
Cyprus	CY	–	–	–	–	–
Czech Republic	CZ	Fuel tax reduction	06/01/22	1.5CZK/l	1.5CZK/l	–
Denmark	DK	–	–	–	–	–
Estonia	EE	–	–	–	–	Control
Finland	FI	–	–	–	–	–
France	FR	Fixed discount	04/01/22	15ct/l	15ct/l	Treatment
Germany	DE	Fuel tax reduction	06/01/22	29.55ct/l	14.04ct/l	Treatment
Greece	GR	–	–	–	–	–
Hungary	HU	Price cap	11/11/21			–
Ireland	IE	Fuel tax reduction	03/10/22	20ct/l	15ct/l	–
Italy	IT	Fuel tax reduction	03/22/22	25ct/l	25ct/l	Treatment
Latvia	LV	–	–	–	–	Control
Lithuania	LT	–	–	–	–	Control
Luxembourg	LU	Fuel tax reduction	03/31/22	7.5ct/l	7.5ct/l	–
Malta	MT	–	–	–	–	–
Netherlands	NL	Fuel tax reduction	04/01/22	17.3ct/l	11.1ct/l	–
Poland	PL	VAT reduction	02/01/22	15%	15%	–
Portugal	PT	"Autovoucher" (limited to 50l/month)	11/01/21	10ct/l	10ct/l	–
Romania	RO	–	–	–	–	–
Slovenia	SI	Price cap	03/15/22	–	–	–
Slovakia	SK	–	–	–	–	–
Spain	ES	Fixed discount	04/01/22	20ct/l	20ct/l	–
Sweden	SE	Fuel tax reduction	06/01/22	17ct/l	17ct/l	–

The retail fuel markets in all countries of our sample are characterized by an oligopoly. Those oligopolists operate nationwide, while there are also smaller suppliers with a single or small number of service stations that operate on a regional basis. For instance, in Germany five firms (Shell, BP/Aral, Esso, Total, and Jet) combine for a market share of 67%. In the other countries, the market shares of the oligopolists are within a comparable range (see Table 2). Differences in the total number of service stations are primarily related to country size and population.

In contrast, the upstream markets in the individual countries of our sample have larger differences. In Austria, for example, there is only one refinery, and the majority of fuel is imported from Germany. France also has a relatively small number of refineries and refining capacity in relation to the market size, resulting in a more inelastic supply side compared to Germany and Italy. Estonia, Latvia, and Lithuania do not have any (or only one) refinery and are therefore also strongly dependent on fuel imports. However, we incorporate these observable differences between countries by including the refinery utilization, imports of crude oil and petroleum products, and the number of gas stations per chain as control variables into our empirical analysis (see Sections 4 and 5).

Considering retail fuel prices, it becomes clear that the price of crude oil accounts for an important share of prices and their fluctuations. Yet, taxes and other duties account for the largest share. Table 3 summarizes the excise duties on gasoline and diesel for the countries in our data set. All countries levy a lower excise duty on diesel than on gasoline, with Germany having the largest diesel privilege (at least without taking into account the temporary fuel tax reductions). Without considering any temporary tax reductions, Austria has the lowest excise duties for fuel and Italy has the highest ones. All countries also levy an additional value-added tax (VAT) on gasoline and diesel.[9] In Germany, an additional fuel carbon tax of 7.2 cents (8.03 cents) on gasoline (diesel) and an additional fuel storage fee of 0.27 cents (0.30 cents) on gasoline (diesel) are levied.

In Germany, the excise duty on fuel ("energy tax") has been lowered by 29.55 cents per liter for gasoline (35.20ct incl. VAT) and by 14.04 cents per liter for diesel (16.70ct incl. VAT) for the period between June 1 and August 31, 2022.[10] With this reduction, Germany has lowered the excise duty on fuel to the minimum level permitted in the EU. In Italy, a reduction of the excise duty on gasoline and diesel by 25 cents per liter (30.50ct incl. VAT) has been introduced from March 22, 2022 on.[11] This measure was initially limited until April 30, but was extended shortly after and ultimately lasted until the end of 2022. The French government has passed a law that introduced a discount for all important fuel products by 15 cents per liter (18.00ct incl. VAT) from April 1, 2022 on.[12] On September 1, the fuel discount has even been increased from 15 to 25 cents per liter and in addition has been extended until December 31, 2022.[13] This discount was paid as a subsidy for quantities sold to the distributor at the second-last distribution level. Based on the termination of the tax reduction in Germany on August 31, 2022 and the simultaneous change of the

Republic, Ireland, the Netherlands, Spain, and Sweden were excluded as additional treated countries for timing reasons. Due to their specific geographical location, data unavailability and/or a currency other than the Euro we decided not to consider Cyprus, Denmark, Finland, Greece, Malta, Romania, and Slovakia.

[9] VAT rates are as follows: 19% in Germany, 20% in Austria, Estonia, and France, 21% in Latvia and Lithuania, and 22% in Italy. To calculate margins and pass-through rates we include VAT reductions associated with the energy tax reductions/discounts to consider the overall reductions.

[10] See: https://www.bundestag.de/dokumente/textarchiv/2022/kw20-de-energiesteuersenkungsgesetz-894664 (Last accessed: July 10, 2023).

[11] See: https://www.gazzettaufficiale.it/eli/id/2022/03/21/22G00032/sg and https://www.loc.gov/item/global-legal-monitor/2022-05-31/italy-new-law-reduces-excise-taxes-and-vat-on-fuels-to-ameliorate-financial-crisis-caused-by-war-in-ukraine/ (Last accessed: July 10, 2023).

[12] See: https://www.legifrance.gouv.fr/download/pdf?id=Yhu9yt95Vn93t_PqwVk79FdAV-pqnhpuwM5LZBeTr90= (Last accessed: July 10, 2023).

[13] See https://www.connexionfrance.com/article/French-news/How-the-French-government-fuel-discount-will-change-from-September-1(Last accessed: July 12, 2023).

Table 2
Overview of relevant market characteristics in all countries considered. When market share values were not publicly available, they were approximated based on the stations in our dataset relative to all stations (denoted by ≈).
Sources: Statistical Report 2023, FuelsEurope, https://www.fuelseurope.eu/publications/publications/statistical-report-2023 (Last accessed: August 7, 2023).

		Austria	Estonia	France	Germany	Italy	Latvia	Lithuania
Downstream	Number of fuel stations	2759	515	11,040	14,452	21,700	600	765
	Oligopoly members	BP, ENI, Jet, OMV, Shell	Alexela Oil, Neste, Circle K, Olerex, Saare Kütus	Shell, Aral, Esso, Total, Jet	Shell, Aral, Esso, Total, Jet	Eni, Q8, Esso, Tamoil	Circle K, Neste, Viada, Virsi-A	Viada, Circle K, Neste, Baltic Petroleum
	Market share of oligopolists	67%	≈54%	62%	67%	49%	≈52%	≈51%
Upstream	Number of refineries	1	0	6	11	10	0	1
	Refinery capacity (in Mt/a)	9.80	0	58.20	100.90	83.90	0	9.60

Table 3
Excise Taxes on Gasoline (E5) and Diesel in cents per liter.
Source: https://ec.europa.eu/taxation_customs/tedb/ (Last accessed: August 7, 2023).

Country	Treatment	Gasoline (E5)				Diesel			
		Pre	Post	Difference	Diff. incl. VAT	Pre	Post	Difference	Diff. incl. VAT
Austria	–	48.00		–	–	40.00		–	–
Estonia	–	42.277		–	–	39.292		–	–
France	04/01/2022	68.29	53.29	−15.00	−18.00	59.40	44.40	−15.00	−18.00
Germany	06/01/2022	65.45	35.90	−29.55	−35.20	47.04	33.00	−14.04	−16.70
Italy	03/22/2022	72.84	47.84	−25.00	−30.50	61.74	36.74	−25.00	−30.50
Latvia	–	41.121		–	–	33.295		–	–
Lithuania	–	43.443		–	–	33.017		–	–

discount in France, we have chosen an observation period until August 31, 2022.

Even though technically the introduced discount in France is different compared to the tax reductions in Germany and Italy, basically it has a similar effect on the costs of the retailers (i.e., service stations). For this reason, it is referred to as a reduction of excise duties paid on the retail-level in the following. In our empirical analysis (see Sections 5 and 6), we compare our estimated coefficients with the overall tax reductions (also including the associated VAT changes), which are also given in Table 3. With regard to the implemented measures, it is important to note that these represented a one-time reduction in all treated countries. In Austria, Estonia, Latvia, and Lithuania, the tax rate remained constant throughout the whole observation period (see Fig. 8 in Appendix C).

4. Data and descriptive statistics

4.1. Data collection

Our analysis is based on several different data sources. First, we scraped data on daily average gasoline (E5) and diesel consumer prices on a service station chain level from the information platform *Fuelo*. These prices on *Fuelo* are the basis of our analysis. *Fuelo* uses official sources as well as information from consumers, publishes this on its website and displays historical information on a daily average level. Real-time price updates are not considered relevant for our analysis as the platform only provides the historical price averages at the service station chain level.[14] The data from *Fuelo* also provides information on the number of fuel stations per service chain. Incorporating this measure serves a dual purpose. Firstly, it helps control for variations in the number of stations across different countries. Secondly, it takes into account the fact that the average fuel prices displayed on the Fuelo website are constructed based on different numbers of chain stations in each country.[15]

Second, we use data on the crude oil price Brent from *Onvista* and exchange rates from Dollar into Euro by the Federal Reserve Bank of St. Louis (FRED) to highlight the relation between consumer prices and the Brent price.[16] The Brent price is also a crucial part to determine retail margins.

Third, we incorporate data on refinery capacities and convert them into a measure of refinery utilization, which indicates how efficiently the refineries are utilizing their maximum capacity. It is crucial to control for refinery utilization in our analysis, since gasoline and diesel can either be produced domestically within the country or imported from other countries. To assess this, we utilize data from both *Concawe* and *Eurostat*.[17] *Concawe* provides annual national-level data on refinery capacities, measured in mega tonnes per annum (Mt/a). With the assistance of *Eurostat* data on the supply (and transformation) of oil and petroleum products, we convert these capacities into a measure of refinery utilization.

Exact calculation of the utilization rate needs some clarification. Before the utilization rate can be determined the domestic production must be calculated from several variables, i.e. the stream of raw oil, loss from refining the crude oil, changes in stock, releases of strategic

[14] Example for German prices from February 2, 2022, https://de.fuelo.net/prices/date/2022-2-2?lang=en. (Last accessed: July 11, 2023). Statement from Fuelo on their sources: https://de.fuelo.net/prices/last_updated?lang=de. (Last accessed: July 11, 2023).

[15] See, Appendix C, Table 9 for the distribution of number of service chain stations in the data. Note: Data from Fuelo has a large market coverage. Example: Germany had 14,452 stations in 2022. Our scraped data covers 10,600 stations.

[16] See historical Brent prices, https://www.onvista.de/rohstoffe/db-Oelpreis-Brent-26262975, and exchange rates from FRED, https://fred.stlouisfed.org/series/DEXUSEU. (Last accessed: July 11, 2023).

[17] See information from Concawe, https://www.concawe.eu/refineries-map/, and Eurostat https://ec.europa.eu/eurostat/databrowser/view/nrg_cb_oilm/default/table?lang=en. (Last accessed: July 11, 2023).

reserve or inflow from marine bunkers.[18] The capacities provided by *Concawe* and *Eurostat* are available on a yearly basis, while the data for refinery utilization is required on a monthly level. To bridge this gap, the yearly capacities are converted into monthly capacities by dividing them by 12. Dividing the monthly domestic production by monthly available capacity determines the utilization of the refinery on a monthly level. Controlling for these refinery utilization possesses the opportunity to rule out differences at the supply side from the local refinery level, i.e. from breakdown in the refinery or loss of access to crude oil.

Fourth, to account for variations in the total imports of oil and petroleum products, we incorporate national-level data from *Eurostat* specifically related to the imports of these products. These imports are measured in thousands of tonnes. By controlling for changes in imports, we aim to capture another aspect of supply-side changes that are likely to be significantly influenced by the outbreak and ongoing war in Ukraine.[19]

4.2. Descriptive statistics

Our final data set includes consumer price data in Euro per liter for the seven European countries Germany, France, Italy, Austria, Lithuania, Estonia and Latvia on a service station chain level during the period from January 3 to August 31, 2022.

Table 8 in Appendix C presents the summary statistics divided by countries. To calculate the margins, we simply subtract taxes and duties as well as the share of the crude oil price (Brent price) attributable to the production of diesel and gasoline from the gross consumer prices.[20] Even though these margins still contain different cost types (e.g., cost of refining, transportation costs), with the crude oil price we can eliminate the main source of input cost variation.

The data set contains 12,515 observations on a service station chain level, i.e. we have a panel data set including price information on 52 unique country-service-station-chain-pairs for 241 days.[21] For each country we observe a different number of chains present in the data (see Table 9 in Appendix C).[22] For instance, Germany has 15 different service station chains present in the data, whereas Austria has six. Overall, there are 30 chains in the treated country's data and 22 chains in the non-treated data.[23] Countries display variations in terms of refinery utilization[24], the number of stations per chain, and total imports of oil and petroleum products. To illustrate this point, a comparison of Germany and Austria serves as an examples. When comparing these two countries, we observe differences in the magnitude of imports of oil and petroleum products (mean: 10,191 v. 987; measured in thousand tonnes), refinery utilization (mean: 0.91 v. 0.55; represented by decimal units) and the number of fuel stations operated per chain (mean: 706 vs. 196). By incorporating these covariates into our analysis, we aim to account for and capture differences in the pre-existing trends and characteristics across countries. These factors help us consider the unique features and dynamics of each country's fuel market.[25]

Despite these differences across countries and service station chains, Table 4 reports that the average price level of diesel and gasoline is very similar across the seven countries, although prices are smaller in Austria, which is mainly driven by the low fuel taxes in this country. Concentrating on the treated countries (France, Germany, Italy), we mostly observe higher average consumer prices after the fuel tax reductions (compare *Pre* and *Post* in Table 4).[26] Even though this seems to be counterintuitive at first glance, this is mainly driven by the increasing price for crude oil during our observation period (see development of the Brent prices in Fig. 1), which has mostly overcompensated the decreased fuel taxes. Table 4 also shows that the absolute as well as the relative retail margins for diesel and gasoline have increased in Italy and Germany after the fuel tax reductions, while they started to decrease in France after the introduction (on average).[27]

Fig. 1 visualizes the development of the average median consumer prices and Fig. 2 shows the daily average retail margins. The figures are divided into sub figures to point out the development of gasoline (upper) and diesel (middle) of the seven European countries during our observation period. The Brent price (lower) is also depicted to highlight the strong link to the market price for crude oil. The vertical lines reflect the introduction of the respective tax reductions in Italy (March 22, yellow), France (April 1, blue), and Germany (June 1, red). In fact, the prices as well as the margins in the seven countries tend to follow the same trend before the policy changes. In all countries, there is also a noticeable increase in both, prices and margins, at the end of February when the war in the Ukraine has started.

With respect to the diesel and gasoline consumer prices, Fig. 1 shows that both have decreased in the first phase after the respective fuel tax cuts in the three treated countries. However, they tend to increase again after a while which is mainly driven by the price increase for crude oil (depicted in dashed gray in the lower sub figure).[28]

Simultaneously, the absolute (and also the relative) effect on retail margins exhibits similar trends between the three treated countries (see

[18] *Eurostat* provides information from the Monthly Oil and Gas questionnaire at page 10, No. 11 on how the gross inland deliveries are determined and we will use this to rearrange this equation for domestic production (https://ec.europa.eu/eurostat/documents/38154/42198/MOS_v2012.1.pdf/f4a7a75c-b0d1-4370-802a-560ca5f86f4d#:~:text=Gross%20inland%20deliveries%20(Observed)%3A,.)%20to%20the%20inland%20market). (Last accessed: July 11, 2023). From the data of *Eurostat* we calculate domestic production for petroleum products with the formula: Domestic Production = Gross inland deliveries − Primary product receipts − Recycled products + Refinery fuel − Imports + Exports + International marine bunkers − Interproduct transfers + Products transferred + Stock changes. Note: Refinery gross output denotes what we call domestic production. Statistics from *Eurostat* regarding what they refer to as *indigenous production* is not available.

[19] See: *Eurostat*, imports of oil and petroleum products by partner country, https://ec.europa.eu/eurostat/databrowser/view/NRG_TI_OILM__custom_6837161/default/table?lang=en. (Last accessed: July 11, 2023).

[20] An important note is that our measure of retail margins includes the refinery margin, the station margin, as well as different cost types such as the cost of refining or the cost of transportation. For a detailed description on the calculation of margins see Appendix B.

[21] The data set is slightly unbalanced.

[22] Fuelo's market coverage per service station chain varies and total market coverage is different across countries. However, it is worth mentioning that the geographic coverage within countries comprises almost their entirety, which can be substantiated through visual inspection. Nevertheless, the goal of this study is to analyze the overall pass-through rate. In this respect, our identification strategy relies on the comparison of the evolution of country-wide large chains average prices, such as most important players Shell, Esso, or Total rather than analyzing the entire market.

[23] The estimation will utilize the not-yet-treated characteristics of the data. During the time periods when part of the data is not yet treated these chains will be used as a comparison group.

[24] See Fig. 9 in Appendix C.

[25] Estonia and Latvia do not have a refinery. For regression purpose the values are set to zero. For E5 in Lithuania (Estonia) there are 9 (2) observations missing which are filled by the last available value of the respective chain to complete the series. For some months in Lithuania refinery utilization is sometimes slightly larger than 1. This probably comes from data accuracy and calculations from an annual capacity to a monthly levels.

[26] It is worth mentioning that the definition of the *Pre* and *Post* periods is distinct for the three treated countries due to the different implementation dates of the fuel tax reductions.

[27] Relative retail margin reflects the simple Lerner-Indices formula, dividing the absolute margins by net prices.

[28] To have a better understanding of the individual consumer price curves, we additionally present the gasoline and diesel price development for the seven countries separately in Fig. 7 of Appendix C.

Table 4
Summary statistics of fuel prices and margins by country before (pre) and after (post) the tax decrease (Numbers in Euro per liter, except the relative margins,). Averages are based on the country and service station chain pairs.

	Country	Austria	Estonia	France		Germany		Italy		Latvia	Lithuania
				Pre	Post	Pre	Post	Pre	Post		
Fuel price	E5	1.783	1.857	1.916	2.038	1.947	1.861	1.948	1.983	1.803	1.738
	Diesel	1.815	1.759	1.886	2.065	1.881	1.968	1.840	1.960	1.753	1.726
Fuel margin	E5	0.273	0.392	0.257	0.392	0.204	0.352	0.228	0.372	0.347	0.270
	Diesel	0.386	0.342	0.323	0.505	0.326	0.465	0.253	0.467	0.387	0.367
Relative fuel margin/Lerner-index	E5	0.260	0.345	0.278	0.330	0.220	0.310	0.260	0.320	0.318	0.262
	Diesel	0.337	0.310	0.324	0.392	0.309	0.374	0.277	0.374	0.338	0.329

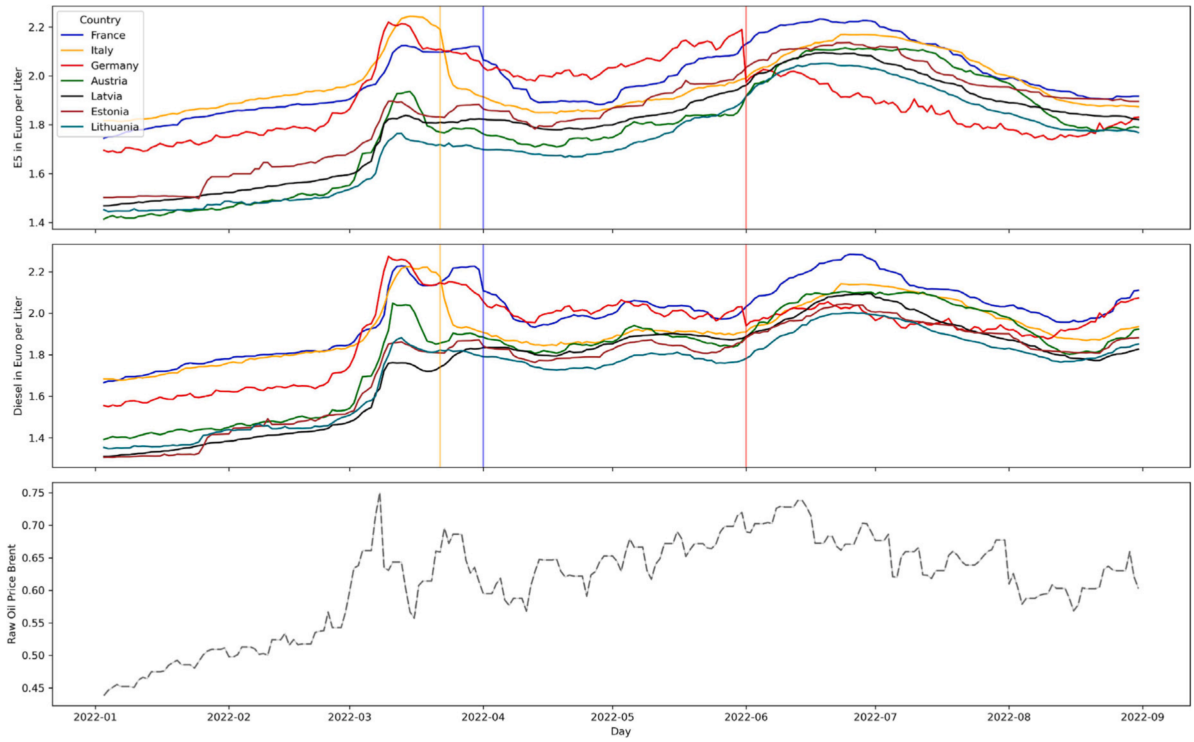

Fig. 1. Development of average consumer prices for gasoline (upper) and diesel (middle). The vertical lines reflect the introduction of the respective tax reductions in Italy (March 22, yellow), France (April 1, blue), and Germany (June 1, red). Brent prices (lower) in Euro per Liter is denoted in dashed gray.

Fig. 2). In addition, the margins reveal a slight difference between the countries. While the margins have increased in Germany for diesel and gasoline, in France and Italy they immediately started to decrease after the temporary tax reduction. In relative terms these margins (see Fig. 10 in Appendix C) reflect the Lerner-Indices (Lerner, 1934; Giocoli, 2012) which range from 0 (no market power) to 1 (monopoly market power). Interpretation of this crude measure of market power is problematic, especially without deep knowledge about the exact cost structure on all parts of the vertical chain within the fuel market and should be carried out with caution (Elzinga and Mills, 2011). Therefore, the large increase and long-term shift in the margins seen in the mid of March 2022 may be prominent but the exact cause cannot be determined without detailed market information on costs and is not part of this paper. In this regard, our empirical analysis will show that these changes in margins are on average not affected by the tax reductions.

5. Methodology

In our empirical analysis, we estimate the impact of the temporary fuel tax reductions on fuel prices and retail margins. In order to do this, we compare the evolution of consumer prices and retail margins at fuel stations in Germany, France, Italy, Austria and the Baltic States, before and after the reductions of the fuel taxes.

We apply a staggered Difference-in-Differences (DiD) design to causally estimate the effect of the temporary fuel tax reductions on fuel prices and retail margins. In contrast to the canonical DiD setup, the staggered design allows to estimate the unbiased average treatment effect on the treated (ATT) when there are more than two time periods and variation in timing of the treatment. This design is more credible and robust than the canonical DiD with a single treatment period because including multiple treatments plausibly alleviates concerns that contemporaneous trends drive the observed treatment effects (see, e.g., Baker et al., 2022). Goodman-Bacon (2021) shows that time-varying treatment effects can create a bias in the static two-way fixed effects (TWFE) DiD estimate since earlier-treated units act as effective controls for later-treated units so that the resultant DiD estimates could reflect differences in treatment effects over time between different treatment groups.

Hence, more recent papers propose alternative DiD estimators that do not suffer from the pitfalls associated with TWFE described above (De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2021; Callaway and Sant'Anna, 2021b). We follow the recent DiD methodology developed by Callaway and Sant'Anna (2021b) as it allows to estimate a time-varying and cohort-specific ATT using not-yet-treated or never-treated as clean controls. Specifically, the estimation strategy

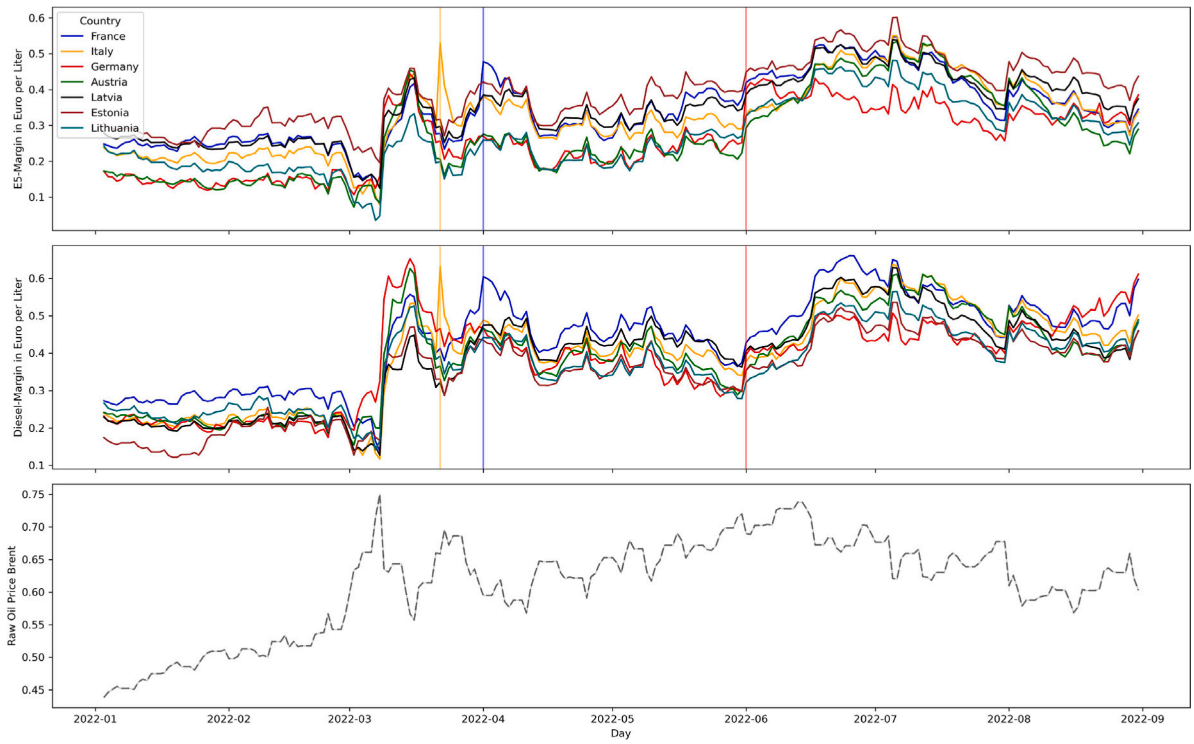

Fig. 2. Development of average retail margins for gasoline (upper) and diesel (middle). The vertical lines reflect the introduction of the respective tax reductions in Italy (March 22, yellow), France (April 1, blue), and Germany (June 1, red). Brent prices (lower) in Euro per Liter is denoted in dashed gray. See relative retail margins in Fig. 10 in Appendix B.

follows the stylized regression[29]:

$$y_{ijt} = X'\beta + \tau_{it} \cdot TAX_{it} + \eta_{ij} + \lambda_t + \epsilon_{ijt}, \tag{1}$$

where y_{ijt} denotes the consumer price (or retail margin) of gasoline or diesel sold by gas station chain j in country i at date t, and TAX_{it} is a dummy variable that equals one when country i implements a temporary fuel tax reduction at date t (note that France, Italy and Germany implemented there reductions at different dates, see Section 3). The vector X' contains our control variables refinery utilization rate, total imports of oil and petroleum imports and number of gas stations.[30] The variable η_{ij} corresponds to country service station chain fixed effects and controls for any time-invariant differences between the countries in our dataset. Finally, λ_t gives the day fixed effects, which capture the transitory shocks that identically affect the individual countries, such as fluctuations in the price of crude oil or the conflict in the Ukraine.

Let us further assume that G_i contains i different states treated at different points of time and C_i is a set of never treated states. Then, under the parallel trend and anticipation assumptions (Wooldridge, 2021) we can estimate the ATT for a treatment-timing group g at a point in time as the group-time average treatment effect using never-treated (2) or not-yet-treated (3) units as controls by using the R package as provided by Callaway and Sant'Anna (2021a)[31]:

$$ATT(g,t) = E[Y_t - Y_{g-1}|G = g] - E[Y_t - Y_{g-1}|C = 1]. \tag{2}$$

$$ATT(g,t) = E[Y_t - Y_{g-1}|G = g] - E[Y_t - Y_{g-1}|D_t = 0, G \neq g]. \tag{3}$$

In the Eqs. (2) and (3), t indexes the time in days, g gives the period in which country i is treated and Y_{it} is the fuel price or retail margin of country i.

Finally, we can average the $ATT(g,t)$ over all countries:

$$\Theta_S(g) = \frac{1}{T - g + 1} \sum_{t=2}^{T} 1\{g \leq t\} ATT(g,t). \tag{4}$$

Eq. (4) then gives the time-average for each group and the overall average respectively. As already mentioned above, we use the fuel prices of seven different European countries to causally identify the effect of the temporary fuel tax reductions. Thereby, Germany, Italy, and France are the treated countries and Austria as well as the Baltic States are the never-treated countries in our staggered DiD approach.

We also want to estimate the treatment effect heterogeneity over time as the effect of the temporary fuel tax reductions on the retail prices might be dynamic. Using an event study design we can prove the process of tax pass-through over time to check whether there is an effect, how many periods it takes to have an effect, and how long it lasts. Moreover, we can test the parallel trend assumption checking the pre-treatment estimators. Hence, based on Eqs. (2) and (3) we provide an event study including pseudo-ATTs for the pre-period and ATTs for the post-period.

To perform the described analysis, various requirements for an unbiased and exogenous estimation have to be satisfied. In general, we can assume that the countries in our data set are very comparable. They are all members of the European Single Market, which implies harmonized border checks, common customs policy, and identical regulatory procedures on the movement of goods within the European Union (EU). Beyond, the seven countries are relatively similar in their geographic location and have highly correlated public and school holidays. In our observation period, also the travel restrictions put in place due to the COVID-19 crisis were similar and no major reforms, which could also affect fuel prices, were implemented.

[29] Depicting this equation's purpose is to intuitively highlight the estimation strategy. Exact estimation will rely on Eqs. (2) and (3) as well as on Footnote 30.

[30] Note that we can add time-invariant variables when using the approach from Callaway and Sant'Anna (2021b) because those variables are interacted with the day fixed effects. Thus, the covariates are not collinear with our state fixed effects, but act more like state-specific time trends. Technically, we use an inverse probability weighting (IPW) to rebalance the distribution of covariates and estimate reweighted ATTs (Abadie, 2005).

[31] See https://bcallaway11.github.io/did/index.html.

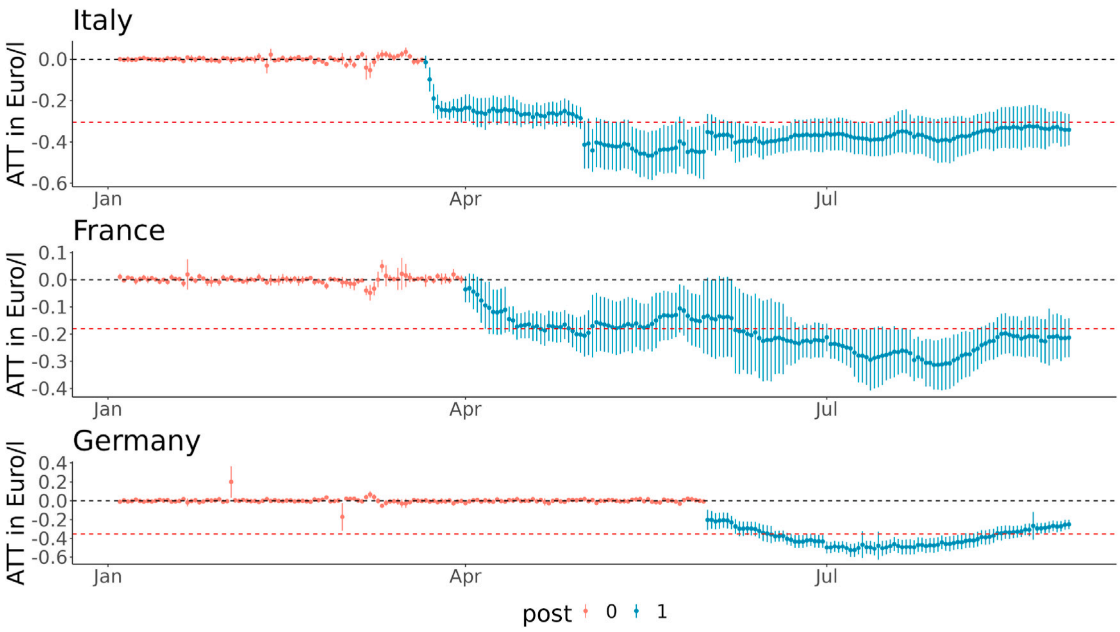

Fig. 3. Event Study of prices with gasoline (E5) and all covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals. Red dashed horizontal line depicts the value for a full-pass-through.

Furthermore, to causally identify an unbiased ATT of the temporary fuel tax reductions on fuel prices, there should also be no other transitory shocks that would differently affect fuel prices in the individual countries before and after the tax reduction. Due to their geographic proximity the petroleum companies in the seven countries procure most of their crude oil from similar sources. Finally, we also focus on a relatively narrow window around the tax reductions, which should alleviate concerns on transitory shocks differently affecting the seven countries.

Moreover, requirements have to be considered in the context of government actions tackling high energy prices observed due to the start of the Ukraine war (see Section 1). The fuel tax reductions were implemented as a measure to counter high inflation rates which may induce the idea of potential endogeneity between tax reductions and price changes, i.e. price changes are affected by the tax reduction and vice versa. There are two major arguments why endogeneity does not pose a problem in our analysis. First, depending on time intervals of tax changes, i.e. tax changes every month, frequent changes might pose a problem for identifying an unbiased effect, especially if long-run relationships are examined (Kaufmann, 2019). However, excise duty reductions by the European governments do not occur on a regular basis,[32] and the interventions in 2022 were implemented at short notice. The energy taxation in Europe is more rigid compared with e.g. the taxation in the US (Kaufmann, 2019) and the interventions in 2022 were a reaction on an exogenous shock, namely the war in the Ukraine. Second, we utilize a panel data set on a service station chain level and a (staggered) DiD design to compare the decisions of countries that implemented a tax reduction to countries that decided against introducing such an intervention. All European countries were equally affected by high fluctuations of the crude oil price (see Fig. 1), which is accounted for by time fixed effects. This design mitigates a bias for the ATT as long as the control groups are not affected by a similar intervention and generally provide a reasonable comparison group to construct an appropriate counterfactual, i.e. the parallel trend assumption must be satisfied (Wooldridge, 2021). Our analysis meets this requirement, especially because the never-treated part of the control countries did not receive any fuel tax reductions during our observation period. Figures presented in Section 6.2 (Figs. 3 and 4) provide evidence in form of event studies, highlighting that the assumption of a common trend can be assumed to be satisfied.

6. Results

6.1. Baseline results

Table 5 presents the results of estimating regression Eq. (4) using the consumer price for gasoline and diesel as outcome variables. The coefficients in columns (I) and (II) correspond to the average treatment effect of the temporary fuel tax reductions on gasoline and diesel in France, Italy and Germany without any other control variables. Columns (III) and (IV) show the effect on consumer prices when we additionally control for the supply side parameters refinery utilization, number of gas stations and the total imports of oil and petroleum products.

In general, the results in Table 5 show that the fuel tax reductions led to a statistically significant decline in the average consumer prices of both fuel types for all three countries treated ($p < 0.001$). For our model without any covariates, in Germany the average price for diesel decreases by 15 cents per liter after the fuel tax reduction (column (II)), whilst the average price for gasoline decreases by about 36 cents per liter (column (I)). Also for France the price decrease for diesel (−18 cents per liter) is slightly lower compared to the one for gasoline (−19 cents per liter). With a price drop of more than 30 cents per liter for diesel and 32 cents per liter for gasoline, also the estimated pass-through rates in Italy are very high. Including additional control variables (columns (III) and (IV)) only quantitatively changes our estimation results, even though we apparently underestimate the average treatment effects without controlling for the supply side.

In a next step, we can calculate the average pass-through rates of the fuel tax reductions. Therefore, we divide the estimated coefficients by

[32] Excise duties on gasoline and diesel have not been changed in Germany since 2006, in Italy since (at least) 2016, and in France since 2018 (prior to the reduction in 2022). For the legal basis, see: https://ec.europa.eu/taxation_customs/tedb/index.html. See also Fig. 8 in Appendix C that highlights exact tax changes in 2022.

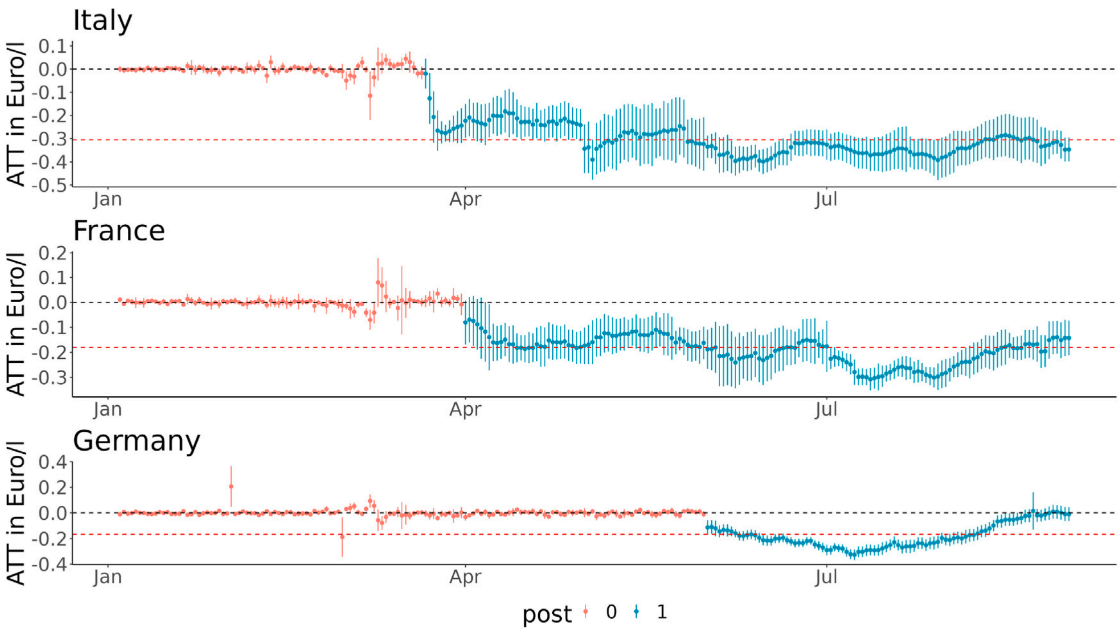

Fig. 4. Event Study of prices with diesel and all covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals. Red dashed horizontal line depicts the value for a full-pass-through.

Table 5
Staggered DiD, referring to estimates of Eqs. (2) and (3) that are averaged by means of Eq. (4). Approach with consumer prices as outcome variable. Bootstrapped (robust) standard errors provided in parentheses are clustered on the country and service station chain level.

	(I) Gasoline	(II) Diesel	(III) Gasoline	(IV) Diesel
Italy	−0.32*** (0.02)	−0.30*** (0.02)	−0.35*** (0.02)	−0.30*** (0.02)
France	−0.19*** (0.02)	−0.18*** (0.01)	−0.20*** (0.02)	−0.19*** (0.01)
Germany	−0.36*** (0.02)	−0.15*** (0.01)	−0.39*** (0.02)	−0.18*** (0.01)
Simple weighted average	−0.31*** (0.01)	−0.20*** (0.01)	−0.33*** (0.01)	−0.21*** (0.01)
Pass-through Italy	106.09%	98.67%	114.27%†	100.00%
Pass-through France	104.65%	99.91%	111.37%	105.99%
Pass-through Germany	103.52%	89.73%	110.32%	106.69%
Time FE	Yes	Yes	Yes	Yes
Country and chain FE	Yes	Yes	Yes	Yes
Supply parameters	No	No	Yes	Yes
Observations	12,515	12,515	12,515	12,515

H0: No effect, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. H0: 100% Pass-Through, † $p < 0.05$, †† $p < 0.01$.

the actual tax reductions in the three countries.[33] The estimated pass-through rates in Table 5 mostly imply a full- or even an over-shifting of the temporary fuel tax reductions. In our baseline estimations (columns (I) and (II) in Table 5), there is an over-shifting for gasoline and almost a full-shifting for diesel. With an estimated pass-through rate of approx. 106%, Italy has the highest passing on of the temporary fuel tax reduction for gasoline and France the highest one for diesel (approx. 100%). Overall, the estimated rates are very similar in the three countries, even though the estimated pass-through rate for diesel is slightly lower in Germany (approx. 90%). Including the control variables for the supply side (columns (III) and (IV) in Table 5) into our regression model generally increases the estimated pass-through rates so that we also find a full- or over-shifting for diesel now. In general, the high pass-through rates might be explained by the inelastic demand for fuel products and particularly by the high public awareness as well as the threat of policymakers to pursue antitrust measures. The 2022 fuel tax reductions had great political and economic implications so that there was a high attention in the public debate (Kahl, 2023).

However, testing whether the average pass-through rates are statistically significant different from a full pass-through (100%) shows that all but one are not different to 100% (see Table 5). Only the pass-through rate of gasoline in Italy in column (III) is statistically significantly higher than 100% ($p < 0.05$).

Beside the general high average pass-through rates, a second interesting finding is that the effects of the tax reductions are mostly higher for gasoline compared to diesel in our estimations. This is in sharp contrast to the literature that finds a more inelastic demand for diesel compared to gasoline (Ajanovic et al., 2012; Karagiannis et al., 2011; Fridstrøm and Østli, 2021). However, the Russian invasion of Ukraine led to a high uncertainty of consumers in the energy markets in 2022, which was combined by an unusually high demand for heating diesel in spring and summer 2022. Households increased their heating diesel stocks out of fear of continuously rising prices, because they expected even higher prices in the future. This phenomena was particularly present in Germany.[34] As heating diesel is a close substitute for diesel (whereas gasoline is not), this might explain the lower pass-through rates for diesel in our results.

[33] For instance, in our baseline estimation for diesel (column (II) of Table 5) the pass-through rate for Germany can be calculated as follows: $passthrough = \frac{EstCoeff}{TaxReduction} = \frac{-15}{-16.7} = 0.8973 = 89.73\%$.

[34] See https://www.dw.com/en/german-residents-make-plans-amid-fears-of-a-winter-gas-shortage/a-62482737.

Table 6
T-test of gasoline versus diesel pass-through rates of Table 5. Depicted values correspond to p-values. See Footnote 35 for the exact construction of the hypothesis test.

	H0: (I) = (II)	H0: (III) = (IV)
Italy	0.4451	0.2392
France	0.7679	0.7403
Germany	0.2582	0.7104

However, scrutinizing differences of the average gasoline estimates against average diesel estimates in each country by means of a hypothesis test reveals no statistically significant differences. Table 6 shows the p-values for each of the tests, with none rejecting the null hypothesis based on common significance levels.[35]

6.2. Pre-treatment trends and dynamic effects

To check whether the estimated results are causal effects and to highlight evolution of pass-through rates over time, we will present an event study design next. The crucial assumption to interpret the results as causally is the parallel trends assumption. Even though this assumption is not directly testable, the event study design does lead to a formal test of pre-treatment trends. With this approach, we can also observe the treatment effects of the fuel tax reductions over time.

Fig. 3 presents the group-time average treatment effects from Eq. (3) for gasoline in the three treated countries.[36] We use the regression model including all control variables and compute bootstrapped 95% confidence intervals. Moreover, we apply a varying base period which means that a pseudo-ATT is computed in each treatment period by comparing the changes in outcomes for a particular group relative to the comparison group in the pre-treatment periods.[37] This just means that we compute changes in the pre-treatment periods from period $t-1$ to period t, but repeatedly change the value of t (Callaway and Sant'Anna, 2021b). The pre-treatment coefficients are close to zero and mostly insignificant in all three countries providing supportive evidence for the common trend assumption. An exception is the time of the beginning of Russia's invasion in the Ukraine, which leads to a short divergence in the pre-trends for Italy and France. However, the pre-trends converge back to the zero line and are statistically insignificant shortly before the exogenous shocks of the tax cuts in all three countries.

In Fig. 3 we can also observe that the treatment effects over time are negative and mostly statistically different from zero. In Germany, there is an immediate drop at the day of the fuel tax reduction with almost full pass-through (red dashed horizontal line). There is a similar development in Italy, where full pass-through is already reached three days after the tax cut. In the following, there is an over-shifting in both countries. On the contrary, in France it takes almost two weeks until there is a full pass-through. This is in line with our theoretical predictions, since France has a more inelastic supply side compared to the two other treated countries (see Section 3). Overall, Fig. 3 suggests that the treatment effects are relatively stable over time in all three countries.

Fig. 4 shows the group-time average treatment effects for diesel from Eq. (3).[38] The pre-treatment coefficients are again close to zero and mostly insignificant (except during the start of the Ukraine conflict for Italy and France). The pattern of the treatment effects over time is very similar compared to the event study for gasoline in Fig. 3. While Italy has a relatively fast full pass-through again, it takes some time in France until there is a significant effect and even longer for a full-pass through. In Germany, we again observe an immediate drop in the treatment effects at the day of the fuel tax reduction. Again, the treatment effects are relatively stable over time, even though the effects get insignificant for Germany in the end of August. This can be explained by the drought in Germany throughout the summer of 2022. The very high temperatures led to exceptional low water levels in German rivers which, in turn, raised the transportation costs for diesel imports (Dovern et al., 2023).

Table 7
Staggered DiD, referring to estimates of Eqs. (2) and (3) that are averaged by means of Eq. (4). Approach with retail margins as outcome variable. Bootstrapped (robust) standard errors provided in parentheses are clustered on the country and service station chain level.

	(I) Gasoline	(II) Diesel	(III) Gasoline	(IV) Diesel
Italy	−0.02 (0.01)	0.00 (0.01)	−0.01 (0.02)	0.02 (0.01)
France	−0.01 (0.02)	0.00 (0.01)	−0.01 (0.02)	−0.00 (0.01)
Germany	−0.01 (0.02)	0.02 (0.01)	−0.02 (0.02)	0.00 (0.01)
Simple weighted average	−0.01 (0.01)	0.01 (0.01)	−0.01 (0.01)	0.00 (0.01)
Time FE	Yes	Yes	Yes	Yes
Country and chain FE	Yes	Yes	Yes	Yes
Supply parameters	No	No	Yes	Yes
Observations	12,515	12,515	12,515	12,515

H0: No effect, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

6.3. Retail margins

Table 7 shows the results of estimating regression Eq. (3) averaged w.r.t Eq. (4) using the retail margins for gasoline and diesel as outcome variables. The results indicate that the reduction in fuel taxes had no significant effect on the average margins in the three countries. This is in line with our results from Section 6.1 as we mostly find a full-shifting of the temporary fuel tax reductions which, on average, should not have an effect on the retail margins.

However, performing an event study design for the outcome variable retail margins implies that there are some positive margins for diesel as well as gasoline in the first days after the fuel tax reductions. Fig. 5 presents the event study results for gasoline in the three treated countries. The margins are significantly positive in the first days after the tax cuts for Italy and France, but insignificantly for Germany. This is in line with our findings in Fig. 3 because it takes some days in Italy and France until the tax reduction is passed through to consumers. Since those daily average treatment effects get insignificantly after a few days, the overall average treatment effects in Table 7 are still

[35] Hypothesis test is constructed by each model and country where the t-statistic is then given by $t = \frac{\frac{EstCoeff_{Gasoline}}{TaxReduction_{Gasoline}} - \frac{EstCoeff_{Diesel}}{TaxReduction_{Diesel}}}{\sqrt{(\frac{StdError_{Gasoline}}{TaxReduction_{Gasoline}})^2 + (\frac{StdError_{Diesel}}{TaxReduction_{Diesel}})^2}}$, following Wooldridge (2015, Chapter. 4-4). Note: The test usually contains a covariance term in the denominator between the two variables. In our comparison the pass-through rates are independent because the underlying estimation is run separately, thus there is no covariance term between gasoline and diesel. Hence, from an econometricians' viewpoint most average pass-through rates are neither statistically different from a complete pass-through nor do the average pass-through rates differ between gasoline and diesel. Overall, the average pass-through rates are mostly in line with findings of the literature on Germany 2022 which imply a full pass-through (Dovern et al., 2023; Schmerer and Hansen, 2023; Kahl, 2023; Bernhardt et al., 2023; Seiler and Stöckmann, 2023).

[36] In Fig. 11 of Appendix C, we present the same dynamic analysis but without including any covariates. The results there are qualitatively very similar.

[37] Pseudo-ATT means that we estimate the effect of participating in the treatment if the treatment had occurred in that period (instead of when it actually occurred).

[38] In Fig. 12 of Appendix C, we present the same dynamic analysis but without including any covariates. The results there are again qualitatively very similar.

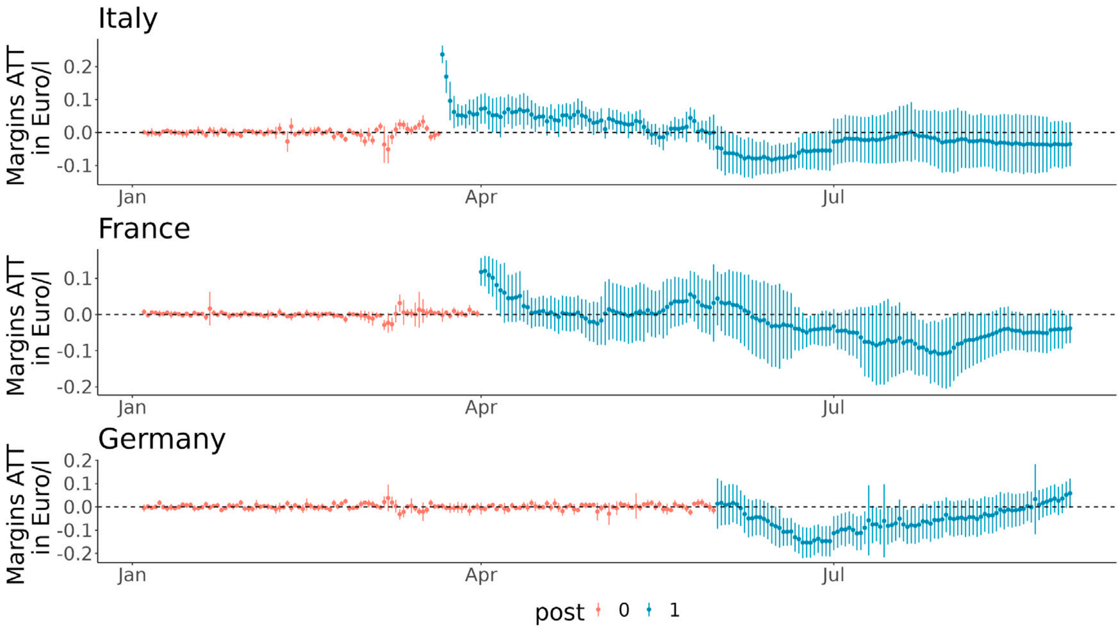

Fig. 5. Event Study of Margins with gasoline (E5) and all covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals.

insignificant. In contrast, there is an immediate drop in Germany in the gasoline prices, which results in the insignificant margins also in the first days after the tax cut there.

In the following weeks, we even find some negative average treatment effects in Fig. 5. For instance, in Germany there is a drop in the estimated ATTs in June and July. Again, this corresponds to our estimated pass-through rates for gasoline (see Fig. 3) because for this months we find an over-shifting of the tax reduction in Germany. Since this means that the German petroleum companies passed through more than 100% of the temporary tax reduction to the consumers at this time, their retail margins are lower compared to the counterfactual scenario where there had been no tax cut.

Fig. 6 shows the equivalent results for the retail margins of diesel. Beside the positive retail margins in Italy and France, we also find some positive diesel margins in the first days after the tax cut for Germany now. During the time of the over-shifting of the fuel tax reductions (see Fig. 4), we even find some significant negative average treatment effects here. This is similar to the estimated retail margins for gasoline (see Fig. 5) and again relates to the fact that the petroleum companies passed through more than 100% of the temporary fuel tax reductions to the consumers at these times leading to lower retail margins compared to the counterfactual scenarios. Overall, those effects over time cancel each other out so that we have no significant average effect in Table 7. This also relates to our main estimations in Table 5 where we find that the temporary tax reductions are mostly passed to consumers on a one-to-one basis on average, which should not lead to any significant changes in the average retail margins.

7. Conclusion and policy implications

This paper provides empirical evidence on the pass-through of temporary fuel tax reductions in the three largest European economies. The governments in Italy, France and Germany introduced relief packages to mitigate the effects of increasing energy prices in the course of post-pandemic economic recovery and the Russian aggression towards the Ukraine. As a part of those packages, the three countries reduced the fuel taxes (introduced a discount on fuel) for several months in 2022. Since the individual measures have taken place at different points of time, we apply a staggered DiD design to causally estimate pass-through rates as well as changes in retail margins.

Our results imply a heterogeneous pass-through over time of the fuel tax reductions depending on the country and type of fuel. Nevertheless, we mostly find a full-shifting of the temporary fuel tax reductions meaning the estimated average pass-through rates are close to 100%. This identifies the fuel markets in the three countries as highly competitive, where the consumers enjoy all of the tax reliefs. High pass-through rates can be explained by the general inelastic demand for fuel products and particularly by the high public awareness as well as the threat of policymakers to pursue antitrust measures during the 2022 tax cuts.

A second finding of our paper is that the average pass-through rates are generally higher for gasoline compared to diesel. However, hypothesis tests indicate that, from a statistical perspective, there is no discernible distinction between the pass-through rates for gasoline and diesel. Besides this statistical perspective, the average findings are in contrast to the previous literature, which finds a more inelastic demand for diesel compared to gasoline. This might be explained by the unusual market situation in 2022. The Russian invasion of Ukraine led to a high uncertainty of consumers in the European energy markets, which (among others) resulted in a higher demand for heating diesel, a close substitute for diesel.

Analyzing dynamics associated with time within the framework of an event study reveals differences with regard to the development of the ATTs between countries and types of fuel. The period of time until a full pass-through is reached for the first time differs, and, in addition, different periods of over- and full-shifting are observable.

With respect to the margins, we find no significant effect different from zero on the average retail margins in the three countries. This is in line with the estimated pass-through rates because the tax reductions were mostly passed on to the consumers one-for-one, which does not change the retail margins of the petroleum companies. However, performing an event study design suggests that the petroleum companies have made some positive retail margins at least in the first days after the fuel tax reductions as it has taken some days until the tax reductions have been fully passed-through to the consumers.

A key takeaway from our paper for policymakers is that temporary fuel tax reductions seem to be a suitable measure to lower consumer

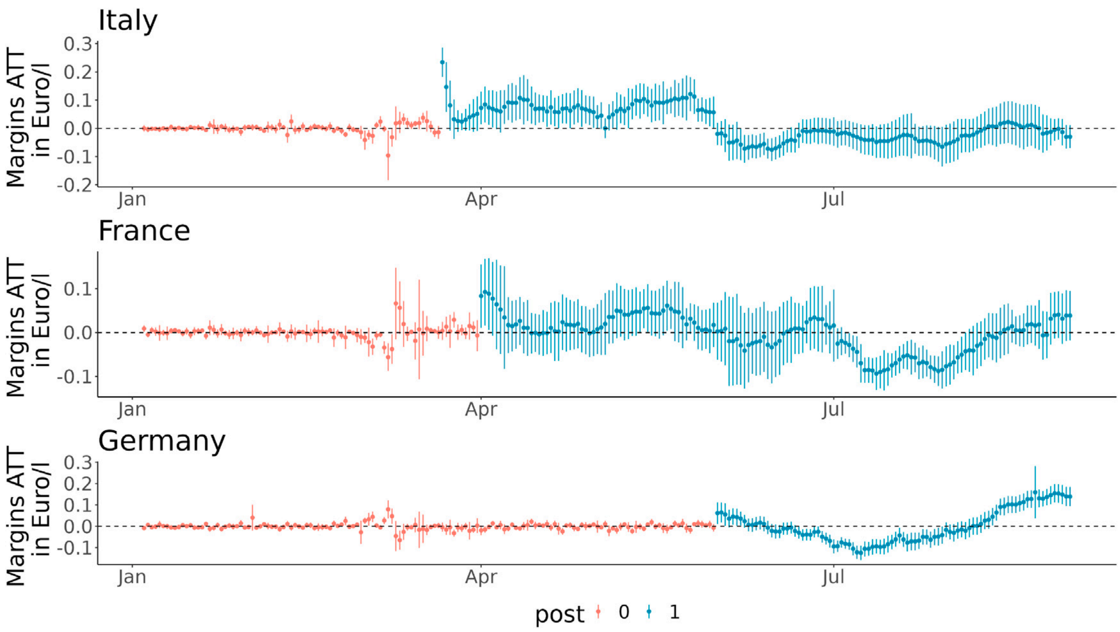

Fig. 6. Event Study of Margins with diesel and all covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals.

prices for diesel and gasoline, even though it may take some time until full pass-through is reached. Hence, the primary goal of the governments to relieve their citizens by achieving lower consumer prices for petroleum products has been met. Whether the corrective goal of a Pigouvian tax or subsidy can be achieved generally depends on whether the consumers also bear the incidence of the measure. In this context, the fuel markets in the three countries seem to be competitive enough so that environmental taxes are passed on to the consumers. However, due to the distributional- and climate-economical shortcomings as well as the relatively high fiscal burden of fuel tax reductions it is debatable whether a temporary fuel tax reduction is an suitable intervention at all.

From a competition policy perspective, our results hardly allow any conclusions to be drawn about whether there are competition restrictions in the fuel market at all. However, the estimated pass-through rates in the three countries imply that the alleged restrictions can at least not hinder a high pass-through of the tax reductions. In general, comprehensive sector analyses by the competition authorities to find the mildest means of competition policy seem to be more appropriate than short-term government interventions in the fuel market.

Apart from already mentioned limitations regarding policy implications, data limitations do not allow to make any statements on welfare effects, as we cannot observe the traded volumes. Furthermore, due to the aggregated price data at service chain level, it is not possible to look at regional effects within individual countries. However, the geographic location of the service stations included in the dataset shows that we observe a balanced geographic coverage, which implies that the average effects within countries are robust. With regard to our observation period and the design of the measures studied, only temporary effects are analyzable. For further studies, it would be interesting to extend the period and also examine the end of the measures and the associated tax increases under the subject of asymmetric pass-through of increasing and decreasing costs, i.e. rockets and feathers. Overall, it is crucial to emphasize that the obtained results are not readily transferable or applicable to other industries. The retail fuel market (and any other market) is characterized by unique features and therefore an own empirical assessment of the pass-through of tax reductions in other industries would be necessary.

Nevertheless, our work provides new and important insights into the transmission of tax reductions in a dynamic and much studied industry, using the most recent methods.

CRediT authorship contribution statement

Chiara Patricia Drolsbach: Software, Validation, Investigation, Data curation, Writing – review & editing. **Maximilian Maurice Gail:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization. **Phil-Adrian Klotz:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We would like to thank Georg Götz, Daniel Herold and Jan Thomas Schäfer for the fruitful discussions. We also thank four anonymous reviewers, participants at the CLEEN 2023 and EARIE 2023 conferences for their comments on preliminary versions of the manuscript. The authors alone are responsible for the content.

Appendix A

Economic theory implies that the elasticities of demand and supply as well as the competitive situation in a market determine the level of pass-through. Following Weyl and Fabinger (2013), we denote p as the retail price and t as the quantity tax rate, so that the pass-through rate is given by $\rho = \frac{dp}{dt}$. We further define the elasticity of demand ($\epsilon_D \equiv -(D'p/Q)$) and supply ($\epsilon_S \equiv S'p/Q$). In this framework, Weyl and Fabinger (2013) postulate that the solution of the firm maximization problem can be described by the conduct parameter $\theta = (p - mc(q))/p\epsilon_D$. θ maps the degree of competition in a market. For instance, θ is equal to 0 in perfect and Bertrand competition, equal to 1 in a monopolistic

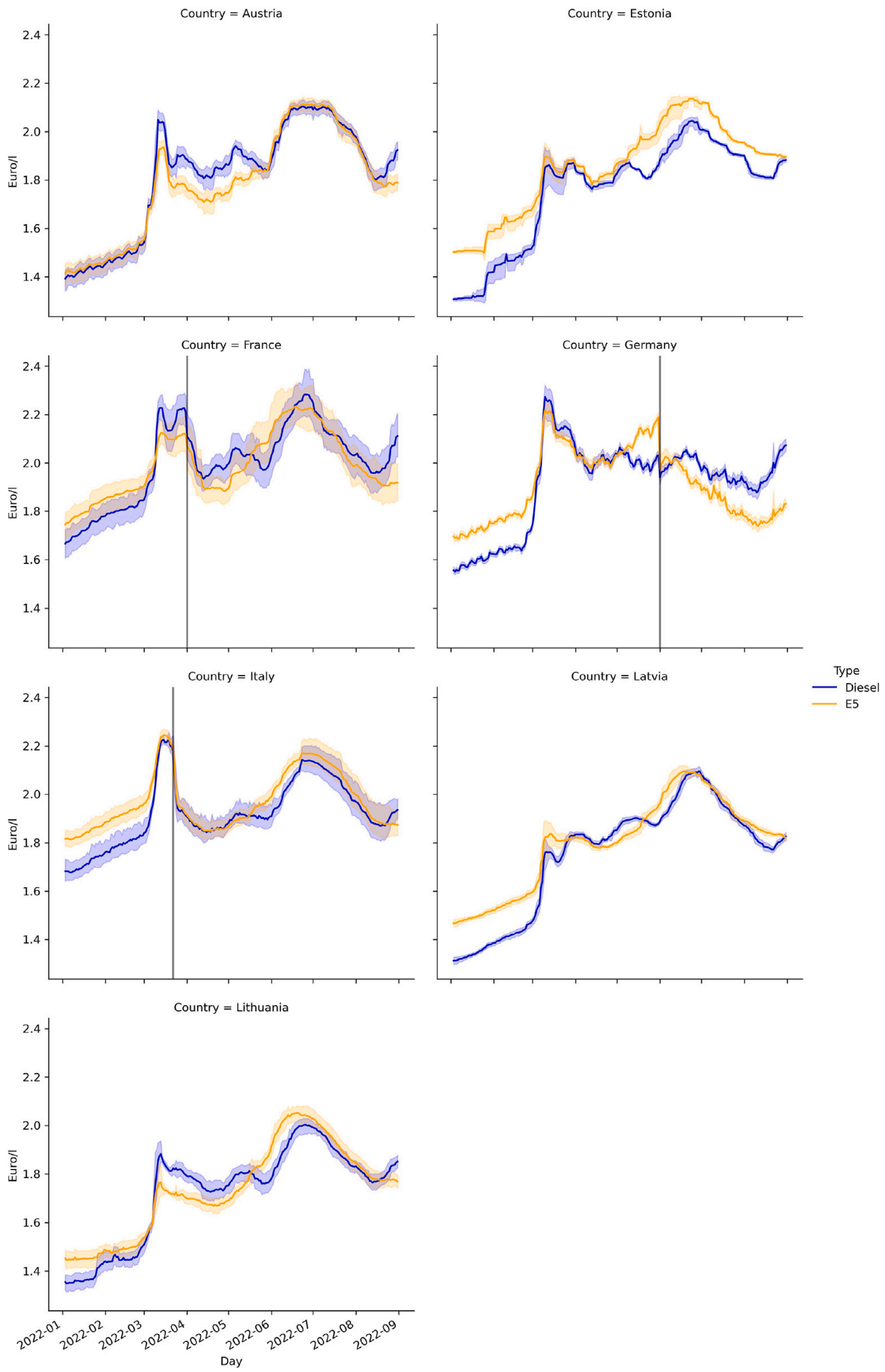

Fig. 7. Development of gasoline and diesel consumer prices for the seven countries in our data set. The vertical lines reflect the introduction of the respective tax reductions. Confidance band is shown to highlight that data varies on the service station chain level.

Table 8
Summary statistics of Austria, France, Germany, Italy, Latvia, Lithuania and Estonia on the service station chain level.

Country	Statistic variable	Count	mean	std	min	25%	50%	75%	max
Austria	Diesel	1446	1.81	0.23	1.28	1.67	1.87	1.99	2.15
	E5	1446	1.78	0.22	1.29	1.61	1.79	1.95	2.16
	Margin of Diesel in Euro/l	1446	0.39	0.13	0.06	0.26	0.40	0.48	0.67
	Margin of E5 in Euro/l	1446	0.27	0.13	−0.01	0.18	0.25	0.37	0.56
	Number of Stations per Chain	1446	196.17	104.28	8.00	117.00	229.50	281.00	312.00
	Relative Margin/Lerner-Index of Diesel	1446	0.34	0.07	0.08	0.29	0.34	0.39	0.50
	Relative Margin/Lerner-Index of E5	1446	0.26	0.08	−0.01	0.21	0.25	0.32	0.43
	Total Imports of Oil and Petroleum Products	1446	987.21	212.86	745.05	779.82	887.39	1, 156.51	1, 328.22
	Utilization of Capacity	1446	0.55	0.27	0.25	0.26	0.37	0.78	0.93
Estonia	Diesel	1190	1.76	0.22	1.26	1.67	1.83	1.90	2.06
	E5	1188	1.86	0.19	1.45	1.74	1.90	1.98	2.16
	Margin of Diesel in Euro/l	1190	0.34	0.11	0.06	0.24	0.37	0.43	0.55
	Margin of E5 in Euro/l	1188	0.39	0.09	0.13	0.32	0.40	0.46	0.61
	Number of Stations per Chain	1190	55.78	30.58	9.00	34.00	59.00	79.00	95.00
	Relative Margin/Lerner-Index of Diesel	1190	0.31	0.07	0.09	0.27	0.32	0.36	0.43
	Relative Margin/Lerner-Index of E5	1188	0.35	0.05	0.13	0.32	0.35	0.38	0.45
	Total Imports of Oil and Petroleum Products	1190	135.61	24.04	103.00	120.00	131.00	147.00	190.00
	Utilization of Capacity	0							
France	Diesel	1687	2.00	0.18	1.52	1.86	2.01	2.13	2.51
	E5	1687	1.99	0.16	1.60	1.88	1.97	2.11	2.50
	Margin of Diesel in Euro/l	1687	0.44	0.14	0.05	0.34	0.45	0.54	0.84
	Margin of E5 in Euro/l	1687	0.34	0.12	0.00	0.25	0.33	0.42	0.74
	Number of Stations per Chain	1687	426.29	216.77	97.00	197.00	419.00	685.00	736.00
	Relative Margin/Lerner-Index of Diesel	1687	0.37	0.07	0.06	0.33	0.37	0.42	0.52
	Relative Margin/Lerner-Index of E5	1687	0.31	0.07	0.00	0.27	0.31	0.36	0.49
	Total Imports of Oil and Petroleum Products	1687	6, 749.68	454.79	6041.00	6, 395.00	6, 965.00	6, 987.00	7, 486.00
	Utilization of Capacity	1687	0.73	0.09	0.60	0.69	0.71	0.85	0.90
Germany	Diesel	3615	1.91	0.19	1.51	1.81	1.97	2.03	2.49
	E5	3615	1.91	0.16	1.65	1.77	1.90	2.04	2.40
	Margin of Diesel in Euro/l	3615	0.38	0.12	0.01	0.26	0.40	0.47	0.96
	Margin of E5 in Euro/l	3615	0.26	0.10	−0.07	0.17	0.26	0.34	0.84
	Number of Stations per Chain	3615	706.67	683.04	13.00	188.00	458.00	980.00	2, 597.00
	Relative Margin/Lerner-Index of Diesel	3615	0.33	0.06	0.01	0.28	0.33	0.38	0.57
	Relative Margin/Lerner-Index of E5	3615	0.25	0.06	−0.09	0.20	0.25	0.31	0.53
	Total Imports of Oil and Petroleum Products	3615	10, 191.52	321.37	9391.28	10, 161.25	10, 305.77	10, 483.69	10, 507.72
	Utilization of Capacity	3615	0.91	0.04	0.84	0.89	0.92	0.95	0.96
Italy	Diesel	1928	1.92	0.15	1.62	1.82	1.89	2.03	2.37
	E5	1928	1.97	0.13	1.75	1.86	1.94	2.07	2.35
	Margin of Diesel in Euro/l	1928	0.40	0.14	0.07	0.31	0.40	0.50	0.80
	Margin of E5 in Euro/l	1928	0.33	0.11	0.04	0.24	0.31	0.39	0.71
	Number of Stations per Chain	1928	2, 109.50	1511.06	176.00	1, 017.50	1, 973.00	3, 054.25	4, 437.00
	Relative Margin/Lerner-Index of Diesel	1928	0.34	0.07	0.07	0.30	0.35	0.39	0.52
	Relative Margin/Lerner-Index of E5	1928	0.30	0.07	0.04	0.26	0.29	0.35	0.49
	Total Imports of Oil and Petroleum Products	1928	6, 446.63	540.48	5701.19	5, 993.60	6, 614.20	6, 986.43	7, 182.83
	Utilization of Capacity	1928	0.86	0.09	0.70	0.82	0.92	0.93	0.97
Latvia	Diesel	1444	1.75	0.23	1.28	1.51	1.82	1.90	2.13
	E5	1444	1.80	0.19	1.44	1.61	1.83	1.93	2.13
	Margin of Diesel in Euro/l	1444	0.39	0.13	0.06	0.24	0.42	0.47	0.64
	Margin of E5 in Euro/l	1444	0.35	0.09	0.03	0.27	0.35	0.41	0.57
	Number of Stations per Chain	1444	64.63	22.20	30.00	41.00	70.00	88.00	89.00
	Relative Margin/Lerner-Index of Diesel	1444	0.34	0.07	0.06	0.28	0.35	0.39	0.47
	Relative Margin/Lerner-Index of E5	1444	0.32	0.05	0.03	0.29	0.32	0.35	0.44
	Total Imports of Oil and Petroleum Products	1444	184.64	49.18	125.43	154.16	180.90	202.69	298.02
	Utilization of Capacity	0							
Lithuania	Diesel	1205	1.73	0.20	1.29	1.55	1.79	1.85	2.04
	E5	1196	1.74	0.19	1.38	1.58	1.74	1.88	2.10
	Margin of Diesel in Euro/l	1205	0.37	0.10	0.10	0.28	0.38	0.45	0.59
	Margin of E5 in Euro/l	1196	0.27	0.10	−0.02	0.19	0.25	0.34	0.53
	Number of Stations per Chain	1205	81.40	25.40	44.00	74.00	76.00	91.00	122.00
	Relative Margin/Lerner-Index of Diesel	1205	0.33	0.06	0.10	0.29	0.33	0.38	0.45
	Relative Margin/Lerner-Index of E5	1196	0.26	0.07	−0.02	0.22	0.26	0.31	0.41
	Total Imports of Oil and Petroleum Products	1205	747.42	245.31	327.60	438.80	764.90	1, 059.60	1, 061.10
	Utilization of Capacity	1205	0.81	0.35	0.20	0.24	1.02	1.06	1.07

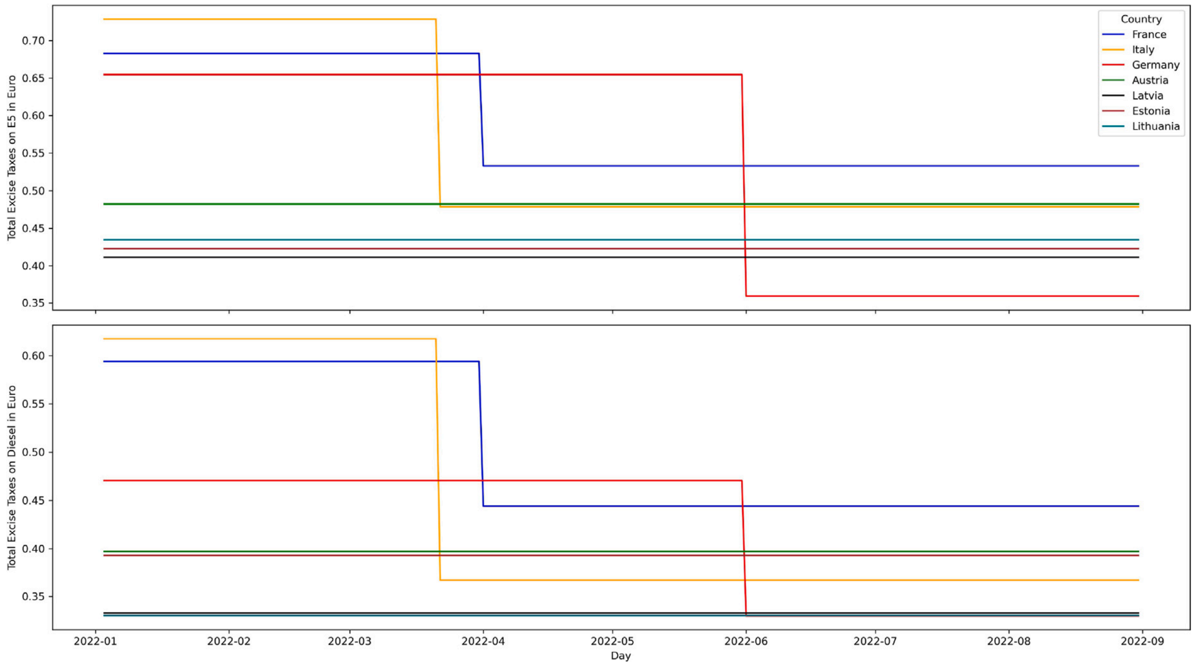

Fig. 8. Development of gasoline and diesel excise duties (inclusive of further duties) for the seven countries in our data set.

market, and equal to $1/n$ in Cournot competition. Then, the pass-through rate ρ is independently of the specific model given by

$$\rho = \frac{1}{1 + \frac{\theta}{\epsilon_\theta} + \frac{\epsilon_D - \theta}{\epsilon_S} + \frac{\theta}{\epsilon_{ms}}}. \tag{A.1}$$

Aside from the conduct parameter θ, formula (A.1) implies that the pass-through of a marginal cost increase also depends on the elasticity of demand ϵ_D, the elasticity of the inverse marginal cost curve (the elasticity of supply)[39] ϵ_S, the curvature of the demand function ϵ_{ms},[40] and the variation of θ in changes of production ϵ_θ.[41]

Even though formula (A.1) suggests that the sign and magnitude of the pass-through is ambiguous, we can simplify the expression for ρ in some special cases. If there is perfect competition in a market ($\theta = 0$), then $\rho = \frac{1}{1+(\epsilon_D/\epsilon_S)}$ so that the pass-through only depends on the ratio of demand and supply elasticity. More generally, if the marginal cost were constant, demand were linear, and θ were constant, expression (A.1) would simplify to $\rho = 1/(1+\theta)$. A rise in the conduct parameter θ (less competition) would lead to lower pass-through in this situation (Genakos and Pagliero, 2022). For instance, in a monopolistic market ($\theta = 1$) the pass-through would be lower ($\rho = 0.5$) compared to a market with perfect competition ($\theta = 0$) where we would have full pass-through ($\rho = 1$).

However, in general, the sign of the effect of an increase in the conduct parameter θ on the pass-through remains ambiguous. This is especially the case for an oligopolistic market, which should be the most appropriate market form to model the fuel industry in Europe. The impact of the conduct parameter on the pass-through can either be positive or negative, depending on the actual market situation. Under certain assumptions also pass-through rates larger than one are possible. Hence, the impact of the intensity of competition on the pass-through rate in an oligopolistic market remains an empirical problem (Genakos and Pagliero, 2022).

Appendix B

To compute the daily average retail margins for the seven countries in our data set, we subtract a fuel share of the crude oil price (major input cost) as well as the country-specific taxes and duties (see Montag et al. (2021)). For each country in our raw data set, we observe a daily average gross consumer price. In a first calculation step, we calculate the average consumer prices without VAT taxes for every day and country (see Footnote 9). To get the daily average net price, we then also subtract the excise duties for the individual countries (see Table 3). Thereby, for the treated countries we have to differentiate between the period before and after the fuel tax reductions.

In a final step, we have to subtract the input cost of crude oil (Brent) from the daily net price. Therefore, we use the information that around 54% of the Brent oil price per barrel corresponds to the production of 19 gallons of gasoline and around 34% to the production of 12 gallons of diesel.[42] We further transform these measures into the input cost per liter of gasoline and diesel. The retail margins of gasoline and diesel are then computed as the average gross consumer price per liter adjusted to VAT taxes and excise duties minus the share of crude oil price per liter of a corresponding fuel product.

[39] The monopolist determines the price based on demand and its costs, there is, just like in an oligopoly, no supply curve and accordingly, no supply elasticity in the sense of perfect competition.

[40] Given by $\epsilon_{ms} = \frac{ms}{ms'q}$, where ms is the negative of the marginal consumer surplus ($ms = -p'q$). If demand is linear, then $\epsilon_{ms} = 1$, if concave, $\epsilon_{ms} < 1$, and if convex, $\epsilon_{ms} > 1$ (and the opposite is also true) (Genakos and Pagliero, 2022).

[41] Given by $\epsilon_\theta = (\theta/q(d\theta/dq))$.

[42] See https://www.eia.gov/energyexplained/oil-and-petroleum-products/refining-crude-oil-inputs-and-outputs.php. (Last accessed: October 19, 2022).

Appendix C

Tables 8 and 9 and Figs. 7–12 provide additional summary statistics, illustration of relevant variables and event studies used in our analysis.

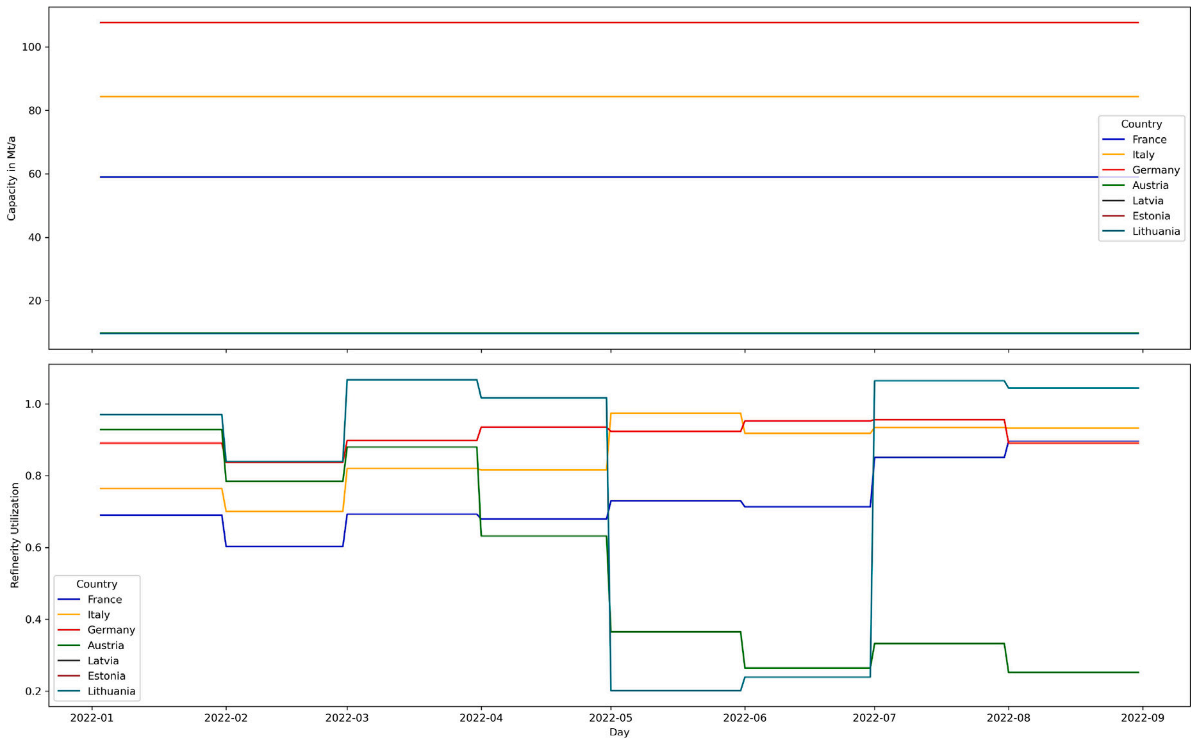

Fig. 9. Development of Capacity and Refinery Utilization.

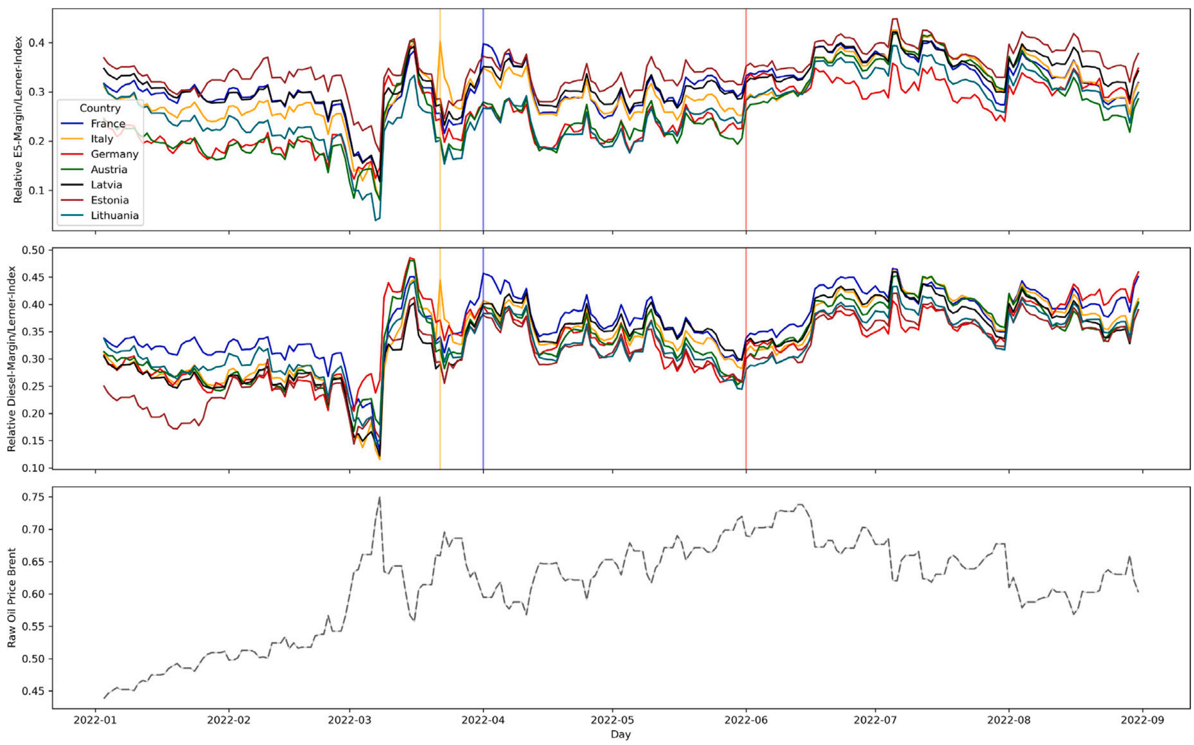

Fig. 10. Development of relative retail margins for gasoline (upper) and diesel (middle). The vertical lines reflect the introduction of the respective tax reductions in Italy (March 22, yellow), France (April 1, blue), and Germany (June 1, red). Brent prices (lower) in Euro per Liter is denoted in dashed gray.

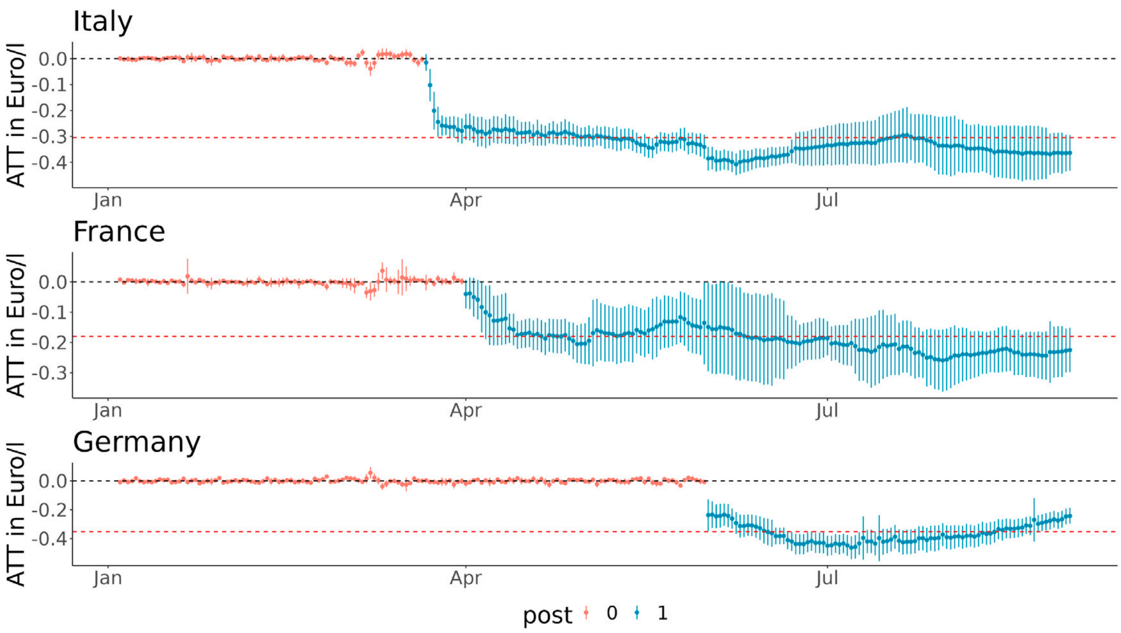

Fig. 11. Event Study of prices with gasoline (E5) and no covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals.

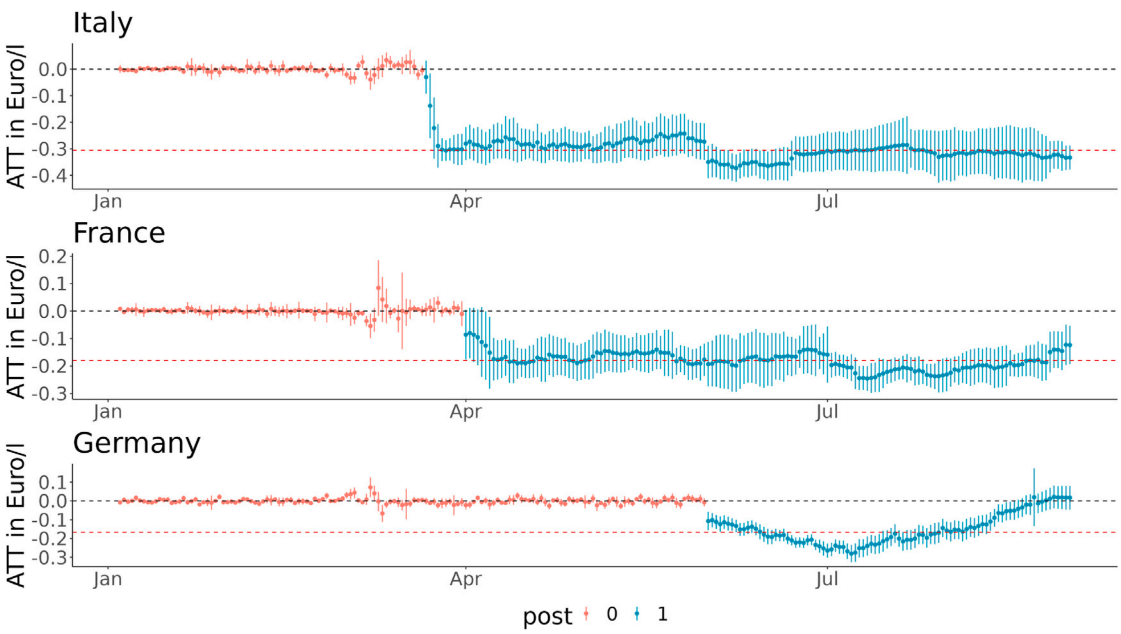

Fig. 12. Event Study of prices with diesel and no covariates. Bootstrapped (robust) standard errors are clustered on the country and service station chain level. Error bars represent 95% confidence intervals.

Table 9
Summary statistics for number of service stations per chain present in the data. Overall coverage: Austria 1177/2748 ≈ 43%, Germany 10600/14500 ≈ 73%, France 2984/11151 ≈ 27%, Italy 16876/21700 ≈ 78%, Estonia 276/491 ≈ 56%, Latvia 388/605 ≈ 64% and Lithuana 407/718 ≈ 56%. Visual inspection of the stations displayed on the map provided by Fuelo reveals the extent of geographical coverage within the national markets.
Source: Fuelo.net, https://de.fuelo.net/gasstations?lang=en.

Country	Provider	Number of stations per chain
Austria	Eni	312
	BP	281
	Shell	248
	OMV	211
	AVIA	117
	Mol	8
Estonia	Terminal	34
	Premium7	9
	NESTE	59
	Circle K	79
	Olerex	95
France	Esso	474
	Shell	97
	Eni	197
	BP	376
	Total	419
	Avia	685
	E.Leclerc	736
Germany	Aral	2597
	Shell	1791
	Esso	1112
	Total	951
	Avia	980
	Gulf	47
	ED	107
	Tamoil	13
	OMV	332
	HEM	383
	agip	437
	Tankpool24	458
	Star	572
	Jet	632
	Westfalen	188
Italy	Esso	2394
	Repsol	176
	IES	206
	TotalErg	1288
	Tamoil	1552
	Q8	2697
	Eni	4437
	IP	4126
Latvia	VIRŠI	70
	Viada	89
	Circle K	88
	NESTE	70
	LN	41
	Astarte	30
Lithuania	Viada	122
	Circle K	91
	Baltic Petroleum	74
	EMSI	44
	NESTE	76

References

Abadie, A., 2005. Semiparametric difference-in-differences estimators. Rev. Econ. Stud. 72 (1), 1–19.
Ajanovic, A., Dahl, C., Schipper, L., 2012. Modelling transport (energy) demand and policies—An introduction. Energy Policy 41 (2012), iii–xiv.
Baker, A.C., Larcker, D.F., Wang, C.C., 2022. How much should we trust staggered difference-in-differences estimates? J. Financ. Econ. 144 (2), 370–395.
Bello, A., Contín-Pilart, I., 2012. Taxes, cost and demand shifters as determinants in the regional gasoline price formation process: Evidence from Spain. Energy Policy 48, 439–448.
Bernhardt, L., Breiderhoff, X., Dewenter, R., 2023. The impact of the tax reduction on fuel prices in Germany–A synthetic difference-in-differences approach. Rev. Econ. 74 (2), 141–160.
Byrne, D.P., De Roos, N., 2019. Learning to coordinate: A study in retail gasoline. Amer. Econ. Rev. 109 (2), 591–619.
Callaway, B., Sant'Anna, P.H., 2021a. Did: Difference in differences. URL https://bcallaway11.github.io/did/, R package version 2.1.2.
Callaway, B., Sant'Anna, P.H., 2021b. Difference-in-differences with multiple time periods. J. Econometrics 225 (2), 200–230.
Clark, R., Assad, S., Ershov, D., Xu, L., 2023. Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market. J. Polit. Econ. forthcoming.
De Chaisemartin, C., d'Haultfoeuille, X., 2020. Two-way fixed effects estimators with heterogeneous treatment effects. Amer. Econ. Rev. 110 (9), 2964–2996.
Dewenter, R., Heimeshoff, U., Lüth, H., 2017. The impact of the market transparency unit for fuels on gasoline prices in Germany. Appl. Econ. Lett. 24 (5), 302–305.
Dovern, J., Frank, J., Glas, A., Müller, L.S., Perico Ortiz, D., 2023. Estimating pass-through rates for the 2022 tax reduction on fuel prices in Germany. Energy Econ..
Doyle Jr., J.J., Samphantharak, K., 2008. $2.00 Gas! Studying the effects of a gas tax moratorium. J. Public Econ. 92 (3–4), 869–884.
Edgeworth, F.Y., 1897. The pure theory of taxation. Econ. J. 7 (25), 46–70.
Elzinga, K.G., Mills, D.E., 2011. The lerner index of monopoly power: origins and uses. Amer. Econ. Rev. 101 (3), 558–564.
Foros, Ø., Steen, F., 2013. Vertical control and price cycles in gasoline retailing. Scand. J. Econ. 115 (3), 640–661.
Fridstrøm, L., Østli, V., 2021. Direct and cross price elasticities of demand for gasoline, diesel, hybrid and battery electric cars: the case of Norway. Eur. Transp. Res. Rev. 13 (1), 1–24.
Fuest, C., Neumeier, F., Stöhlker, D., 2020. The Pass-Through of Temporary VAT Rate Cuts: Evidence from German Retail Prices. Technical Report. ifo Working Paper.
Fuest, C., Neumeier, F., Stöhlker, D., 2022. Der Tankrabatt: Haben die Mineralölkonzerne die Steuersenkung an die Kunden weitergegeben? Perspekt. Wirtsch. 23 (2), 74–80.
Genakos, C., Pagliero, M., 2022. Competition and pass-through: evidence from isolated markets. Am. Econ. J.: Appl. Econ. 14 (4), 35–57.
Giocoli, N., 2012. Who invented the lerner index? Luigi amoroso, the dominant firm model, and the measurement of market power. Rev. Ind. Organ. 41, 181–191.
Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. J. Econometrics 225 (2), 254–277.
Grasso, M., Manera, M., 2007. Asymmetric error correction models for the oil–gasoline price relationship. Energy Policy 35 (1), 156–177.
Harju, J., Kosonen, T., Laukkanen, M., Palanne, K., 2022. The heterogeneous incidence of fuel carbon taxes: Evidence from station-level data. J. Environ. Econ. Manage. 112, 102607.
Kahl, M.P., 2023. Was the German Fuel Discount Passed on to Consumers?. Leuphana Universität Lüneburg, Institut für Volkswirtschaftslehre.
Karagiannis, S., Panagopoulos, Y., Vlamis, P., 2011. Symmetric or asymmetric interest rate adjustments? Evidence from Southeastern Europe. Rev. Dev. Econ. 15 (2), 370–385.
Kaufmann, R.K., 2019. Pass-through of motor gasoline taxes: Efficiency and efficacy of environmental taxes. Energy Policy 125, 207–215.
Kaufmann, R.K., Laskowski, C., 2005. Causes for an asymmetric relation between the price of crude oil and refined petroleum products. Energy Policy 33 (12), 1587–1596.
Lerner, A.P., 1934. The concept of monopoly and the measurement of monopoly power. Rev. Econom. Stud. 1 (3), 157–175, URL: http://www.jstor.org/stable/2967480.
Li, S., Linn, J., Muehlegger, E., 2014. Gasoline taxes and consumer behavior. Am. Econ. J.: Econ. Policy 6 (4), 302–342.
Marion, J., Muehlegger, E., 2011. Fuel tax incidence and supply conditions. J. Public Econ. 95 (9–10), 1202–1212.
Maskin, E., Tirole, J., 1988. A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles. Econometrica 571–599.
Montag, F., Sagimuldina, A., Schnitzer, M., 2021. Does tax policy work when consumers have imperfect price information? theory and evidence.
Noel, M., 2009. Do retail gasoline prices respond asymmetrically to cost shocks? The influence of Edgeworth cycles. Rand J. Econ. 40 (3), 582–595.
Noel, M.D., 2015. Do Edgeworth price cycles lead to higher or lower prices? Int. J. Ind. Organ. 42, 81–93.
Noel, M.D., et al., 2011. Edgeworth price cycles. In: New Palgrave Dictionary of Economics. Palgrave Macmillan.
Perdiguero-Garcia, J., 2013. Symmetric or asymmetric oil prices? A meta-analysis approach. Energy Policy 57, 389–397.
Schmerer, H.-J., Hansen, J., 2023. Pass-through effects of a temporary tax rebate on german fuel prices. Econom. Lett. 227, 111104.
Seiler, V., Stöckmann, N., 2023. The impact of the German fuel discount on prices at the petrol pump. Ger. Econ. Rev. 24 (2), 191–206.
Silvia, L., Taylor, C.T., 2014. Tax pass-through in gasoline and diesel fuel: the 2003 washington state nickel funding package increase.
Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J. Econometrics 225 (2), 175–199.
Weyl, E.G., Fabinger, M., 2013. Pass-through as an economic tool: Principles of incidence under imperfect competition. J. Polit. Econ. 121 (3), 528–583.
Wooldridge, J.M., 2015. Introductory Econometrics: A Modern Approach. Cengage learning.
Wooldridge, J., 2021. Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. Available at SSRN 3906345.

5 Efficiency in COVID-19 Vaccination Campaigns - A Comparison across Germany's Federal States

Authors: Georg Götz, Daniel Herold, Phil-Adrian Klotz and Jan Thomas Schäfer

Published in: *Vaccines 2021, 9, 788.*

vaccines

Article

Efficiency in COVID-19 Vaccination Campaigns—A Comparison across Germany's Federal States

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

Department of Economics, Justus Liebig University Giessen, Licher Strasse 62, 35394 Giessen, Germany; georg.goetz@wirtschaft.uni-giessen.de (G.G.); phil.a.klotz@wirtschaft.uni-giessen.de (P.-A.K.); jan.t.schaefer@wirtschaft.uni-giessen.de (J.T.S.)

* Correspondence: daniel.herold@wirtschaft.uni-giessen.de

Abstract: Vaccination programs are considered a central pillar of the efforts to stop COVID-19. However, vaccine doses are scarce and several organizational and logistical obstacles, such as the timing of and reserves for second shots and delivery failures, apparently slow down vaccination roll-outs in several countries. Moreover, it is an open question as to where vaccines are administered as efficiently as possible (vaccination centers, hospitals, doctor's offices, pharmacists, etc.). The first aim of our study was to systematically evaluate the efficiency of a country's vaccination campaign. The second aim was to analyze how the integration of doctors' offices into a campaign that formerly relied only on vaccination centers affected the speed of that campaign. Using data on vaccine deliveries and vaccinations given in Germany, we find considerable differences across federal states in terms of efficiency, defined as the ability to administer the most vaccinations out of a given number of available doses. Back-of-the-envelope calculations for January to May 2021 show that vaccinations would have been 3.4–6.9% higher if all federal states had adopted a similar ratio between vaccinations given and vaccines stored, as the most efficient states did. This corresponds to 1.7–3.3% of Germany's total population. In terms of our second research goal, we find evidence that the integration of doctors' offices into the vaccination campaign significantly increased the ratio of vaccinations administered out of a given stock of vaccine doses. On average, there appears to be a structural break in this ratio after doctors' offices were integrated into the vaccination campaign on 5 April 2021. On average, an additional 11.6 out of 100 available doses were administered each week compared to the period prior to that date. We conclude that there are considerable regional differences in the efficiency of the vaccination roll-out. Systematic efficiency analyses are one step to detecting inefficiencies and to identify best practices that can be adopted to eventually speed up the vaccination roll-out in a country.

Keywords: COVID-19; vaccination campaign; efficiency

Citation: Götz, G.; Herold, D.; Klotz, P.-A.; Schäfer, J.T. Efficiency in COVID-19 Vaccination Campaigns—A Comparison across Germany's Federal States. *Vaccines* **2021**, *9*, 788. https://doi.org/10.3390/vaccines9070788

Academic Editor: Barbara Rath

Received: 27 May 2021
Accepted: 9 July 2021
Published: 14 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

On 31 December 2019, the World Health Organization was informed of a new kind of infectious disease that was first identified in Wuhan, China [1]. It took only a few months for the virus, later known as SARS-CoV-2, to become a pandemic and spread across the globe [2,3]. By the time this paper was written, the total number of casualties was over 3.9 million worldwide, with over 180 million confirmed cases (data provided by WHO as of 28 June 2021, see https://covid19.who.int/, accessed on 28 June 2021).

Vaccination is considered a central pillar in the effort to stop the COVID-19 pandemic [4]. Despite the scarcity of vaccines, and subject to a prioritization of vulnerable and/or exposed individuals within the population [5,6], the policy goal is clear: roll out the available doses as quickly and efficiently as possible. However, the administrative and logistical challenges are substantial [7–9]. For instance, many countries report difficulties in the logistics of the vaccination roll-out, including unexpected delivery failures and the timing of second shots [10]. To overcome these obstacles, vaccine reserves are built. When reserves are too low, appointments have to be re-scheduled and no more shots can be given. When reserves are too high, more vaccinations could be given without compromising second shots. Vaccinations are at a sub-optimal level in both situations. It is important to detect and eventually avoid these inefficiencies that prolong the pandemic and cost lives [11–13]. This article analyzes this problem using Germany and its federal states as an example.

Another important aspect of the vaccination roll-out's efficiency is the question of where the population can receive vaccinations. Potential candidates include vaccination centers, hospitals, retail pharmacies or doctors' offices. The integration of general practitioners into a country's vaccination campaign has been especially discussed in the literature, for instance, in terms of overcoming vaccine "hesitancy" in the population [14].

The vaccination roll-out in Germany started in December 2020. Even though every federal state faces essentially the same problem of determining the optimal level of vaccine reserves to maintain a smooth vaccination campaign, the 16 federal states show noticeable differences in the progress of their respective vaccination roll-outs. The vaccination campaign in Germany was based in vaccination centers until 5 April 2021, when doctors' offices were officially integrated into the campaign (https://bit.ly/35PaZoi, accessed on 28 June 2021).

The aims of our study are two-fold. The first aim is to determine a measure of the efficiency of the federal states' vaccination campaigns. Data Envelopment Analysis (DEA) is used to analyze relative efficiency by systematically comparing vaccine deliveries and stocks (input) with vaccinations given (output). Those federal states that are able to maintain a smooth vaccination campaign with the lowest vaccine reserves are identified as efficient. Efficiency is interpreted in relative terms, thus the most efficient federal states constitute a lower bound on efficiency. The benefits of decreasing inefficiencies can be approximated in a counterfactual scenario where it is assumed that every federal state adopts the ratio between vaccinations given to available doses as the most efficient states. The second aim of our study is to investigate the effect the integration of doctors' offices had on Germany's vaccination campaign. We do so using a fixed effects panel regression.

2. Materials and Methods

2.1. Dataset & Ethical Approval

Our dataset is comprised of two sources of data. Data on vaccine deliveries are available on the website of the Federal Ministry of Health, https://bit.ly/3vQAu3a (accessed on 28 June 2021). Data on daily vaccinations are published online by the Robert Koch Institute (RKI), https://impfdashboard.de/ (accessed on 28 June 2021). On the official website of the RKI, only recent data on vaccinations are available. However, historical data are made available via github.com by members of the German public broadcaster ARD, https://bit.ly/3vWPzQH (accessed on 28 June 2021). Our observation period was 27 December 2020 (when the first delivery arrived) to 16 May 2021.

The data used in this study are publicly available, highly aggregated and completely anonymized. We therefore consider this study exempt from ethical review.

In Germany, two mRNA-based vaccines produced by Biontech/Pfizer and Moderna and one vector-based vaccine produced by Astra-Zeneca were used. Towards the end of the observation period, a second vector-based vaccine produced by Johnson&Johnson was approved. According to contemporary guidelines in Germany, immunization with mRNA-based vaccines (and the vaccine produced by Astra-Zeneca) required two shots that had to be given within 6 weeks (12 weeks). Only one dose of the vaccine produced by Johnson & Johnson was required. Official guidelines were provided by the Federal Institute of Vaccines and Biomedicines (Paul Ehrlich Institute) (https://bit.ly/2Qud8le, accessed on 28 June 2021); an overview of the relevant information on storage requirements can be found, for example, here: https://bit.ly/2RmHy9J (accessed on 28 June 2021).

The data published by RKI are frequently revised ex post. Occasionally, RKI reported daily vaccinations of zero for some federal states. Missing vaccinations are apparently attributed to subsequent days by RKI. As a consequence, the data presented here might differ slightly from the aggregated figures published by RKI.

2.2. Method: DEA

A DEA is a method for comparing the relative efficiencies of different Decision Making Units (DMUs). The method was formalized by [15]. Since then, it has been used to analyze the performance of water [16] or electricity [17] suppliers as well as railroad firms [18]. In the health care sector, the method has been applied to hospitals [19,20] and also to vaccination centers [21]. The concept is closely related to cost-effectiveness and cost-utility analyses in health economics [22,23]. For an extensive review of the various applications and refinements of the technique, see [24].

In a DEA, efficiency is measured as a deterministic ratio between inputs and outputs. In the original formulation of [15], there are $j = 1, \ldots, n$ DMUs with $r = 1, \ldots, s$ outputs and $k = 1, \ldots, m$ inputs. The known values of output r and input k of DMU j are denoted by y_{rj} and x_{kj}, respectively. To find the most efficient among the n DMUs, the following program can be used:

$$\begin{aligned} \max_{u_r, v_k} \quad & h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{k=1}^{m} v_k x_{k0}} \\ \text{s.t.} \quad & \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{k=1}^{m} v_k x_{kj}} \leq 1 \; \forall j = 1, \ldots, n, \\ & v_r, u_k \geq 0 \; \forall r = 1, \ldots, s, k = 1, \ldots, m. \end{aligned} \tag{1}$$

The weights v_r, u_k are endogenously determined by a comparison of all DMUs included as a reference. The problem can then be reformulated to yield a program that is solvable via linear programming. A detailed derivation can be found in [15] or in textbooks such as [25].

In addition to a DEA with constant returns to scale (CRS), DEAs were computed under the assumption of variable returns to scale (VRS). In a VRS DEA, the production possibility frontier is non-linear and defined by multiple DMUs. The concept of VRS DEA goes back to [26] and is sometimes also referred to as the BCC model. Note that the BCC model has a slightly different optimization problem; for a more detailed overview, see Chapter 2 in [25]. The question of whether CRS or VRS is assumed is especially important in applications in the health care sector [25], Ch. 16.4.4.2, especially when it comes to the analyses of vaccination centers [21].

The DEA is an appropriate method for analyzing the relative efficiency levels of the vaccination roll-out of Germany's federal states for the following reasons. The DEA is a non-parametric method, that is, no assumptions on the functional form of the production function have to be imposed. Second, the DEA is used to analyze the relative performance of not-for-profit entities. In the case at hand, the overall policy goal is to maximize output, that is, to roll out as many vaccinations as possible given the available doses. This is in contrast to, for example, a profit maximizing firm that takes into account the price effects of its output choice.

Three types of models were computed with different output variables. In models T, 1S and 2S, the respective output variables are the total number of shots given, the total number of first shots given and the total number of second shots given in week t. The observation period for model 2S starts on 17 January 2021, the day the first second shot was recorded. Comparing the results of models 1S and 2S allows for an identification of the prioritization of federal states. A federal state that has high scores in model 2S but performs relatively poorly in model 1S can be considered to prioritize the full immunization of the population.

The input variable is the sum of vaccine deliveries in week t and vaccine reserves in week $t-1$. This variable approximates the amount of doses available for vaccination in week t. Even though this variable potentially overestimates the number of available doses in absolute terms (e.g., because doses arriving towards the end of week t might not be available for vaccination in that week), it does not systematically bias the comparison of the federal states because every state is affected in the same way (see the fluctuations in deliveries presented in Section 3.1 The results are robust to variations in the input variable. As an example, the results of DEAs computed with vaccine reserves in $t-1$ are presented in Appendix A.

In this study, separate DEAs were computed for each week. A federal state can then be considered efficient if it receives high scores in many periods. This was done because the scores of a single DEA might be biased, for example, due to large deliveries to a federal state (or the lack thereof) towards the end of the respective observation period, which can lead to relatively high (or low) inputs in relation to outputs compared to federal states that did not (or did) receive larger shipments. A further advantage of examining the results of multiple DEAs is that the results become more robust against outliers resulting from, for example, errors in the data or public holidays in some but not all federal states.

2.3. Method: Counterfactual Scenario

To illustrate the potential impact of improvements in efficiency on the progress of the vaccination campaign, a back-of-the-envelope calculation was carried out. A ratio of total vaccinations given in week t (output) to vaccine deliveries in week t and vaccine reserves at the end of week $t-1$ (input) was determined. Formally, federal state i's share in week t reads:

$$s_{i,t} = \frac{\text{vaccinations}_{i,t}}{\text{deliveries}_{i,t} + \text{reserves}_{i,t-1}}. \tag{2}$$

For example, $s_{i,t} = 0.6$ would indicate that in federal state i in week t, 60% of the doses available for vaccination in week t are actually administered, whereas the remaining 40% are held back as reserves.

The following counterfactual scenario was assumed to compute the potential gain in vaccinations administered when efficiency is improved. Suppose federal states k and l are identified as the most and second most efficient federal states, respectively, by the DEA described in Section 2.2. In the counterfactual analysis, it was assumed that each federal state that is identified as inefficient by the DEA adopts the ratio of vaccinations given to available doses in k in each week t. Formally, $s_{i,t} = s_{k,t}$ for all i and t. A more conservative perspective was taken by assuming that each federal state, except for k, adopts the ratio of the second most efficient federal state l, $s_{i,t} = s_{l,t}$ for all $i \neq k$ and t.

2.4. Method: Vaccination at Doctor's Offices

Based on the data, it was tested whether there was a structural break in $s_{i,t}$, as defined in Equation (2), after 5 April 2021, when general practitioners were integrated into Germany's vaccination campaign. In doing so, the following fixed effects panel regression was estimated:

$$s_{i,t} = \alpha + \beta t + \delta D_{\text{April 5}} + D_i + \epsilon_{i,t}. \tag{3}$$

The dependant variable in Equation (3) is the share $s_{i,t}$. Note that this is the ratio between (unweighted) output and input of DEA model T (see above). The observed shares are depicted in Section 3.1 for each federal state.

In Equation (3), $s_{i,t}$ is explained by a time trend t, the variable $D_{\text{April 5}}$ takes the value 1 for the period after 5 April 2021 and 0 otherwise, as well as a federal state specific fixed effect D_i. The latter controls for time-invariant effects specific to a federal state. This approach allows for an investigation of whether the integration of doctors' offices into the German vaccination campaign potentially improved the efficiency of the campaign on average.

3. Results

3.1. Results: Descriptive Statistics

In Figure 1, daily vaccine deliveries and vaccines administered are depicted for each federal state as well as for Germany as a whole (DE) for the period from 27 December 2020 to 16 May 2021. One can see that deliveries are relatively infrequent and their magnitude varies. This corroborates the organizational challenge of the vaccination campaign, especially against the background that it is necessary to administer two shots to achieve full immunization. Remarkably, Saarland (SL) received a delivery of over 80,000 vaccines at the end of March. These deliveries, however, seem to have barely affected contemporary vaccinations. Given that Saarland continued to receive relatively high deliveries in subsequent weeks, this indicates that reserves built up. Recall that observations of zero or even negative vaccinations administered stem from errors in the data, which are corrected by RKI ex post.

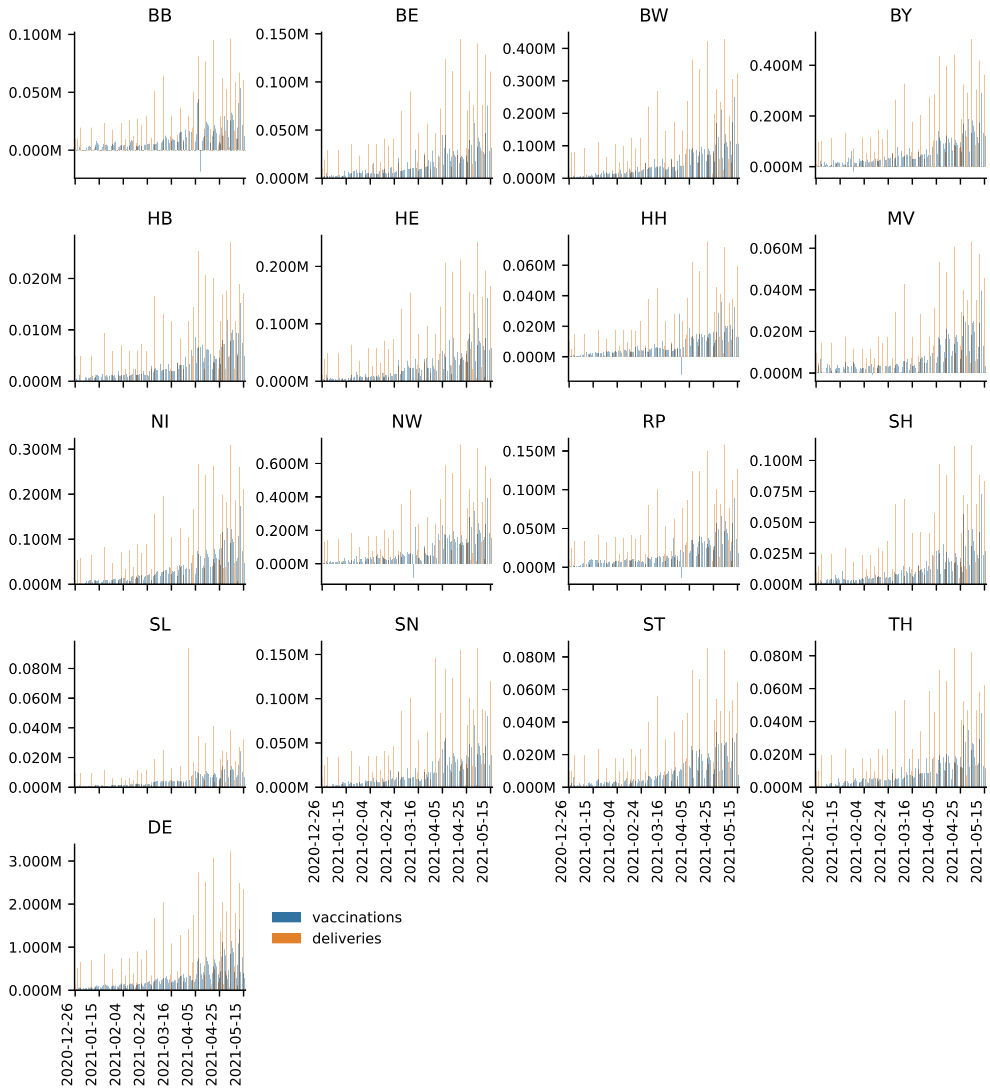

Figure 1. Daily number of people vaccinated and deliveries by Federal State.

Table 1 presents an overview of total deliveries, vaccinations broken down by first and second shots, and storage quotas for the first quarter of 2021. The storage quota relates deliveries to total vaccinations. For instance, Bremen (HB) had 6164 doses in stock, which was approximately 4.66% of total deliveries. Note that Table 1 provides a snapshot: storage quotas are inflated in some federal states (e.g., Saarland (SL); see above) that received large shipments by the end of the first quarter whereas others did not.

Table 1. Summary statistics for q1.2021 by Federal State.

Federal State	Deliveries	Total Vaccinated	First	Full	Storage Quota
HB	132,255	126,091	90,234	35,857	4.66%
TH	432,150	403,299	280,682	122,617	6.68%
BE	697,200	635,103	427,101	208,002	8.91%
BB	472,590	429,998	328,564	101,434	9.01%
SH	550,575	494,930	378,776	116,154	10.11%
NI	1,503,825	1,328,867	920,694	408,173	11.63%
DE	15,974,175	14,005,686	9,763,805	4,241,881	12.32%
HH	347,325	301,837	211,080	90,757	13.10%
BY	2,569,635	2,232,881	1,516,433	716,448	13.11%
BW	2,085,075	1,798,007	1,253,952	544,055	13.77%
HE	1,188,420	1,022,298	691,217	331,081	13.98%
RP	790,185	676,868	495,688	181,180	14.34%
ST	418,200	351,265	251,865	99,400	16.01%
NW	3,363,300	2,739,983	1,913,865	826,118	18.53%
MV	300,225	234,839	159,250	75,589	21.78%
SN	856,875	654,525	405,999	248,526	23.61%
SL	266,340	176,290	129,287	47,003	33.81%

Figure 2 presents vaccine reserves broken down by federal states for the period from 27 December to 16 May 2021. For Germany as a whole (DE), towards the end of the observation period there were almost 6 million doses in stock. This constitutes jabs for over 7% of Germany's population (83,190,556 people as of 30 September 2020, based on information provided by the Federal Statistical Office, https://bit.ly/3bOABoR, accessed on 28 June 2021). One can see a sawtooth-like shape of deliveries and vaccinations given; however, the level of reserves drastically increases over time. Even though some federal states, such as Bremen (HB), apparently reduced vaccine reserves to some degree during the observation period towards mid May, every federal state seems to have built up substantial vaccine reserves. Again, daily data have to be interpreted with care due to the ex-post revisions of RKI.

Finally, Table 2 presents descriptive statistics on reserves per first doses given. For instance, in Brandenburg, an average of 0.47 doses were held as reserves per first shot given with a median of 0.39. Consistent with the observations described above, Bremen shows relatively low median reserves whereas federal states, such as Saarland (SL) and Lower-Saxony (NI), apparently had relatively high median reserves.

Table 2. Summary statistics on reserves per first dose given for the period from 11 January 2021 to 16 May by Federal State. The first two weeks of the campaign were left out here due to the low number of vaccinations in relation to deliveries.

Index	#Weeks	Mean	Std	Min	25%	50%	75%	Max
BB	18	0.47	0.22	0.24	0.31	0.39	0.58	0.88
BE	18	0.35	0.14	0.16	0.25	0.32	0.43	0.68
BW	18	0.56	0.32	0.2	0.28	0.5	0.76	1.34
BY	18	0.35	0.12	0.19	0.24	0.33	0.41	0.55
DE	18	0.42	0.17	0.2	0.24	0.42	0.51	0.7
HB	18	0.28	0.16	0.03	0.13	0.29	0.41	0.55
HE	18	0.56	0.28	0.21	0.3	0.51	0.78	1.05
HH	18	0.39	0.16	0.2	0.22	0.4	0.43	0.82
MV	18	0.36	0.21	0.07	0.18	0.26	0.56	0.68
NI	18	0.62	0.34	0.19	0.28	0.59	0.92	1.12
NW	18	0.54	0.3	0.14	0.19	0.59	0.75	1.08
RP	18	0.31	0.11	0.17	0.21	0.27	0.41	0.51
SH	18	0.31	0.12	0.15	0.23	0.26	0.4	0.56
SL	18	0.58	0.29	0.21	0.34	0.53	0.81	1.2
SN	18	0.58	0.25	0.28	0.35	0.55	0.74	1.12
ST	18	0.52	0.24	0.21	0.26	0.55	0.69	0.9
TH	18	0.46	0.35	0.13	0.23	0.38	0.51	1.55

Figure 3 illustrates that $s_{i,t}$ has substantial fluctuations over time in most federal states and differs remarkably between them. For instance, in Bremen (HB), that share relatively quickly increases to values over 0.4 and even jumps to over 0.8, whereas in Saarland (SL) the share never exceeds 0.5.

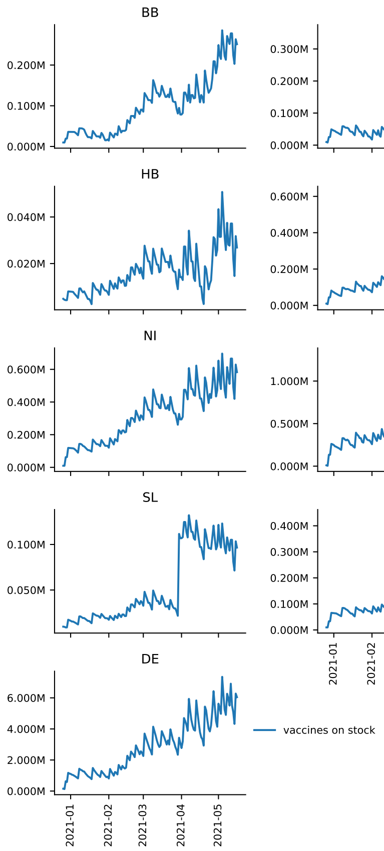

Figure 2. Daily number of vaccines in stock by Federal State.

3.2. Results: DEA

Table 3 presents the results of the mean DEA scores in Models T, 1S and 2S of DEAs performed for each week of our observation period. Here, constant returns to scale (CRS) were assumed.

Table 3 shows that Bremen was assigned the highest average DEA score with 0.8289 in model T, 0.7842 in model 1S and 0.689 in model 2S. This indicates that Bremen (HB) had the most efficient vaccination roll-out in Germany under the CRS assumption.

One might argue that the results of a small federal state, such as Bremen with a total population of less than 600,000, are not applicable to larger federal states such as Bavaria (BY) or Northrhine-Westphalia (NW) with populations of 13.8 million and almost 18 million, respectively, and a lower population density. Likewise, Bremen receives fewer vaccine deliveries (see above), which potentially eases the organizational burden of the vaccination roll-out. In other words, a vaccination campaign might be susceptible to decreasing returns to scale so that it becomes increasingly difficult to distribute vaccinations the larger the input of vaccines. Thus, the scores of VRS DEAs for each week of the observation period were computed. The average scores are presented in Table 4.

The results presented in Table 4 show that Bremen is assigned average DEA scores of 1 in every model. This means that Bremen is among the federal states that define the production possibility for every week. In contrast to the CRS DEA, larger federal states, such as North Rhine-Westphalia, were assigned significantly higher DEA scores.

Potential drivers of the different levels of efficiency are explained in Section 4. Appendix A contains a more thorough discussion of the results of the DEA as well as a robustness check.

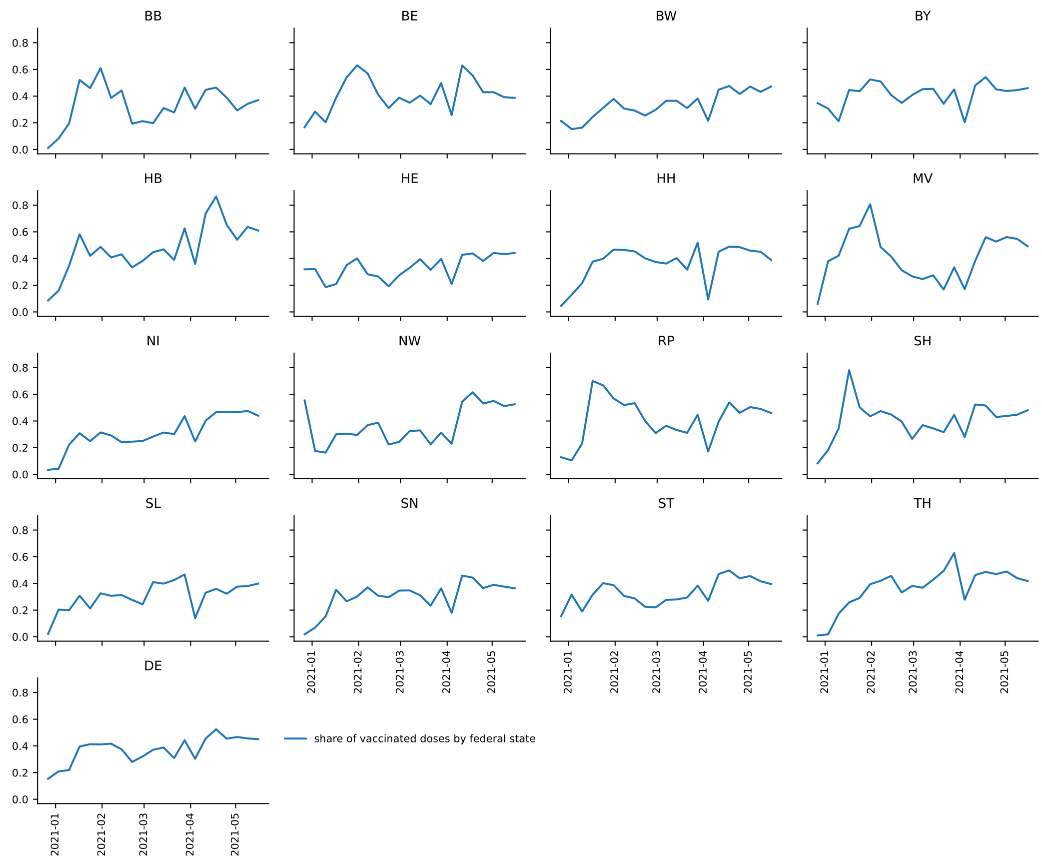

Figure 3. The share of vaccinations given in week t in relation to deliveries in week t and reserves in week $t-1$ $\left(s_{i,t} = \frac{\text{vaccination}_{i,t}}{\text{deliveries}_{i,t}+\text{reserves}_{i,t-1}}\right)$ for all federal states.

Table 3. Average efficiency scores for the period from 27 December 2020 to 16 May 2021 for DEAs performed on a weekly basis under the CRS assumption.

Federal State	T	1S	2S
BB	0.5699	0.5248	0.4321
BE	0.7195	0.6708	0.6857
BW	0.5891	0.584	0.4581
BY	0.7371	0.7276	0.5477
HB	0.8289	0.7842	0.689
HE	0.5993	0.597	0.4873
HH	0.6495	0.643	0.4942
MV	0.7165	0.6591	0.4911
NI	0.5442	0.5483	0.3971
NW	0.6428	0.6554	0.4438
RP	0.7113	0.6455	0.5263
SH	0.7125	0.6531	0.5363
SL	0.5527	0.5643	0.4078
SN	0.5321	0.4814	0.5457
ST	0.5898	0.5712	0.4695
TH	0.6544	0.6155	0.6566

Table 4. Weekly average DEA efficiency scores for the period from 27 December 2020 to 16 May 2021 under the VRS assumption.

Federal State	T	1S	2S
BB	0.6487	0.6077	0.5159
BE	0.8042	0.7375	0.7332
BW	0.7346	0.7397	0.6222
BY	0.9408	0.9067	0.7495
HB	1	1	1
HE	0.6994	0.6813	0.5605
HH	0.7185	0.7187	0.5909
MV	0.7887	0.7602	0.6215
NI	0.682	0.6771	0.4849
NW	0.9793	0.9347	0.8258
RP	0.8004	0.7317	0.6078
SH	0.7873	0.7404	0.6273
SL	0.6446	0.6662	0.5935
SN	0.6071	0.5433	0.5984
ST	0.6499	0.6329	0.5401
TH	0.7277	0.6807	0.7274

The results presented in Tables 3 and 4 indicate differences in the prioritization of first and second shots between federal states. Apparently, federal states with lower scores in model 1S than in model 2S (e.g., Saxony, SN) focus on full immunization of the population whereas federal states with lower scores in model 2S than in model 1S (e.g., Lower Saxony, NI) seem to prioritize first shots.

3.3. Result: Counterfactual Scenario

Assuming that all federal states adopt the ratio (2) for Bremen ($s_{\mathrm{HB},t}$) or North Rhine-Westphalia ($s_{\mathrm{NW},t}$) in each week t, we compute hypothetical vaccinations given per federal state. The results are presented in Table 5.

Table 5. Counterfactual scenario where it is assumed that less efficient federal states adopt $s_{i,t}$ for i = Bremen or i = North Rhine-Westphalia every week t.

		Bremen		Northrhine-Westphalia	
Federal State	**Act. Vacc**	**Hyp. Vacc**	**%-Gain**	**Hyp. Vacc**	**%-Gain**
BB	1,114,386	1,273,349	14.26%	1,232,573	10.61%
BE	1,703,655	1,881,650	10.45%	1,819,474	6.80%
BW	5,268,410	5,630,827	6.88%	5,449,360	3.43%
BY	6,368,489	6,813,909	6.99%	6,594,566	3.55%
HB	357,057	357,057	0	357,057	0%
HE	2,967,320	3,213,995	8.31%	3,110,320	4.82%
HH	854,467	927,992	8.60%	898,988	5.21%
MV	794,509	825,545	3.91%	798,212	0.47%
NI	3,817,325	4,089,342	7.13%	3,955,702	3.62%
NW	8,869,664	9,171,453	3.40%	8,869,664	0%
RP	1,970,369	2,086,120	5.87%	2,018,964	2.47%
SH	1,396,932	1,486,886	6.44%	1,438,825	3.00%
SL	523,240	582,246	11.28%	565,857	8.14%
SN	1,912,294	2,144,315	12.13%	2,079,466	8.74%
ST	1,024,732	1,120,043	9.30%	1,084,781	5.86%
TH	1,041,879	1,120,897	7.58%	1,086,081	4.24%
Total	39,984,728	42,725,627	6.85%	41,359,888	3.44%

The results presented in Table 5 indicate that over 2.7 million (+6.85%) more vaccinations would have been given until 16 May 2021 if all federal states had adopted $s_{\mathrm{HB},t}$ in each week t. This corresponds to 3.29% of the German population.

A comparison with North Rhine-Westphalia, whose vaccination roll-out is remarkably efficient given the size of the federal state, yielded less optimistic, yet noticeable results. According to the figures presented in Table 5, almost 1.4 million more doses would have been administered if all federal states (except for Bremen) had adopted North Rhine-Westphalia's ratio between vaccinations given and reserves and deliveries. This still corresponds to 1.65% of the entire population and constitutes a plus of around 3.44%.

3.4. Results: Vaccination at Doctor's Offices

The results of the fixed effects panel regression Equation (3) can be found in Table 6.

Table 6. Output table for Equation (3).

	Dep. Var. $s_{i,t}$	
$D_{\text{April 5}}$	0.116 **	(4.08)
Time trend	0.000941	(0.35)
Constant	0.333 ***	(13.43)
Observations	320	
R^2	0.348	
R^2 adjusted	0.311	

t statistics in parentheses; two and three asterisks correspond to $p < 0.001$ and $p < 0.01$, respectively.

The results presented in Table 6 indicate that there was a statistically significant structural break on 5 April 2021, for Germany. That is, $s_{i,t}$ increased by 11.6% on average in the period after 5 April 2021, compared to the period prior to that date. In other words, the share of doses held back as reserves decreased by the same fraction on average. This means that, on average, 11.6 more out of 100 available doses—measured by vaccine reserves plus vaccine deliveries—were given each week.

The results indicate that the structural break diagnosed on average for Germany seems to be driven by some federal states that exhibit a relatively strong structural break (e.g., NW). Not every federal state shows a statistically significant structural break in the period after 5 April 2021. Moreover, on average, based on the results presented in Table 6, no statistically significant time trend was diagnosed. These findings are supported by the various robustness checks presented in Appendix B. These findings not only capture the effect of the integration of doctors' offices into the vaccination campaign as is discussed in Section 4.

4. Discussion

The primary aim of the present study was to determine a lower bound on efficiency in the German federal states' vaccination campaigns. Several DEAs were performed and Bremen apparently defines the efficiency frontier during the observation period. Among the larger federal states, North Rhine-Westphalia receives remarkably high efficiency scores. With VRS, this federal state has relatively high average efficiency scores, whereas under CRS its average scores are relatively low.

A counterfactual scenario was computed based on a back-of-the-envelope calculation. It was shown that an increase in efficiency could have led to an increase in vaccinations in the magnitude of 3.44–6.85%, which corresponds to 1.65–3.29% of the German population. This shows that analyses of the efficiency of vaccination roll-outs can play an integral role in overcoming the COVID-19 pandemic. Avoiding excessive reserves is crucial—a vaccine that is unused cannot save lives. That countries handle the vaccine doses available to them as efficiently as possible seems to be particularly important against the background of pronounced vaccine scarcities in low-income countries (e.g., [27]).

By the time this paper was written, 53% (34.5%) of the German population had received a first (second) shot. In comparison to the other 30 countries of the EU/EEA, Germany ranks eighth (sixth) when it comes to first (second) shots. The German vaccination campaign was slower compared to some non-EU countries: the United States (first shots: 53.5%, second shots: 45.5%), Canada (67.4%, 25.6%), Israel (64%, 59.6%) and the United Kingdom (65%, 47%) (see https://ourworldindata.org/covid-vaccinations, accessed on 28 June 2021). In the literature, organizational and country-specific factors that influence the speed and efficiency of the COVID-19 vaccination roll-out are identified. Potential drivers of the high speed of Israel's vaccination campaign include the small size of the country both in terms of area and population, a relatively young population, an efficient health care system with IT-heavy organization, large vaccine orders and a clear prioritization system for vaccinations within the population in the early phases of the distribution process [28,29]. In particular, Israel relaxed its prioritization system at some point to avoid diminishing returns [30]. A similar strategy was pursued in the USA [31]. There is also evidence that the use of online communication effectively increased the speed of the United States' vaccination campaign [31,32]. In contrast to that, for example in the UK, appointments were initially allocated mostly by text messages or mobile phone calls. The vaccination campaign in the UK was sped up when officials decided—as one of the first countries worldwide—to extend the interval between first and second shots of two important vaccines in order to vaccinate as many people as possible at least once [33].

Given the little information that is publicly available regarding the administrative processes of the vaccination roll-out in the different federal states, it is difficult to pinpoint certain aspects where the federal states' organization of the vaccination roll-out differs. Possible explanations for the observed differences in efficiency include diverging practices when it comes to building reserves to, for example, avoid the re-scheduling of future appointments (https://bit.ly/2RoBAoO; all links in this section accessed on 28 June 2021), the handling of appointments for the different priority groups (https://bit.ly/3ob1Leu), the use of more efficient syringes (https://bit.ly/3hyfDyj) and the administration of appointments (https://bit.ly/3tFoi4h).

According to the results presented in Section 3.2, Bremen's vaccination campaign is apparently the most efficient. It is remarkable that the state's vaccination campaign was not solely planned and executed by Bremen's government. Local firms have supported the campaign by establishing a vaccination initiative ("Bremen impft", https://bit.ly/3x1JmVd). This collaboration between private firms and public officials is responsible for the administration of the—by the contemporary standards as of March 2021—largest vaccination center in Germany (https://bit.ly/3gWP05b). Moreover, it is documented that there was close cooperation with health insurance companies to systematically identify and allocate appointments to high-risk patients in Bremen (https://bit.ly/3y1Bi6Y).

The second aim of this study was to analyze the effect that the integration of general practitioners into Germany's vaccination campaign on 5 April 2021 had on efficiency. The results indicate that there was a structural break for Germany as a whole in the period after that integration. On average, 11.6 more out of 100 available doses were vaccinated per week compared to the period before 5 April. This indicates that the integration of doctors' offices into the campaign has sped up the vaccination roll-out.

While in other countries, such as the UK, general practitioners have been part of the vaccination campaign from the beginning (https://nyti.ms/2UOTi6h), in Germany, doctors' offices were officially integrated into the campaign roughly four months after the beginning of the vaccination roll-out. Approximately 35,000 general practitioners started to administer COVID-19 vaccinations in the week after Easter in Germany. With around 102,000 doctors' offices in Germany, this is around one third (https://bit.ly/35Y3748). In total, practitioners have ordered 1.4 million doses for vaccination in calendar week 14 (https://bit.ly/3x2yLcw). This number has continuously increased in subsequent weeks. Based on the dataset used in this paper, deliveries to doctors' offices were around 2.5 million doses in calendar week 19, on a par with those to vaccination centers.

General practitioners order the desired number of vaccine doses for week $t + 1$ until Tuesday of week t at local pharmacies. The orders of each federal state are centrally monitored to ensure a fair allocation of doses across Germany. Physicians are informed on Thursday of week t about how many doses they will receive in week $t + 1$. Appointments are scheduled at the general practitioner's discretion (see https://bit.ly/3qEHPSB). Practitioners were officially constrained by the prioritization system until 7 June 2021, when the system was abandoned (see https://bit.ly/3qxJPMs). Administering COVID-19 vaccinations is associated with a relatively large bureaucratic burden on general physicians, which is considered especially cumbersome because those vaccinations are an extra service offered by practitioners [34].

The results of the present paper have implications for clinical management. First, expectation management seems to be important. An efficient vaccination roll-out with as little reserves as possible can make it necessary to communicate that appointments (especially for first shots) might be re-scheduled in case of delivery disruptions. Second, integrating doctors' offices into a vaccination campaign can apparently speed up a vaccination roll-out that is otherwise based on vaccination centers. Third, the results above indicate that the more efficient management of appointments (reserves for second shots, (re-)scheduling, identification of vulnerable individuals) can be considered a way to improve efficiency. The German vaccination campaign appears to be characterized by a high degree of bureaucracy. Even though further research is necessary, more flexible, innovative and IT-based solutions can be expected to speed up the vaccination roll-out.

The present study has some limitations. First, it constitutes a rather high level approach. A lower bound on efficiency, rather than an optimal inventory management, was determined. The latter could be analyzed, for example, in an (s,S)-model [35]. Such an approach requires a more sophisticated dataset including, for instance, information about planned vaccine deliveries. Moreover, the decision makers' expectations have to be accounted for and the trade-off between first and second shots has to be discussed [36]. Second, the study at hand suffers from a lack of information about the administrative details and the causes of the differences in detected efficiency levels. The DEAs could also be enriched by more detailed information about health care workers, the number of vaccination centers, demographic and geographic factors, and so forth. Third, based on the data available to the authors by the time this paper was written it was impossible to analyze whether the vaccination of the most vulnerable individuals in terms of their risk of mortality was actually prioritized. Fourth, systematic, historical data on vaccine deliveries to doctors' offices and to vaccination centers per federal state were also not available to the authors. This study therefore relied on a statistical analysis of the time series of vaccinations given and the available doses. As such, the structural break we identified in Section 3.2 might not only be driven by doctor's offices being integrated into the vaccination campaign. It cannot be excluded that other factors also affected $s_{i,t}$ (see Equation (2)) and influenced the results. For instance, one of these factors could be the findings of [37] that a 12-week, rather than a shorter, timespan between first and second shots of the vaccine produced by Astra-Zeneca does not reduce protection against COVID-19. This might have led decision makers to reduce stock holdings of Astra-Zeneca's vaccine, that is, all else being equal, leading to an increase in $s_{i,t}$. If this happened (with some delay) in the period after 5 April 2021, the effect of the integration of doctors' offices into the vaccination campaign would be overestimated.

Most of the above limitations stem from a lack of appropriate data. If these data become available in the future, further analyses on the topic will be fruitful.

5. Conclusions

This study is a first attempt to systematically analyze the efficiency of the COVID-19 vaccination roll-out in different regions of a country. This exercise allows one to identify a lower bound of efficiency of that country's vaccination campaign. Similar analyses can be performed for other countries as well, especially because data requirements are minimal.

We used Germany as an example. Our findings indicate that efficiency comparisons, such as DEA, can be valuable for detect ing inefficiencies in a vaccination roll-out. Our results on the effect of integrating general practitioners into the vaccination campaign indicate an important avenue for how the administration of vaccinations might be sped up.

Author Contributions: Conceptualization, G.G. and D.H.; methodology, J.T.S., P.-A.K. and D.H.; software, J.T.S. and P.-A.K.; validation, G.G., D.H., J.T.S. and P.-A.K.; formal analyses, G.G., D.H.; investigation, G.G., D.H.; data curation, J.T.S.; writing—review and editing, G.G. and D.H.; visualization, J.T.S. and P.-A.K.; supervision, G.G. and D.H.; project administration, D.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study, because publicly available, highly aggregated and anonymized data were used.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in our study are publicly available on the website of the Federal Ministry of Health (https://www.bundesgesundheitsministerium.de/coronavirus/faq-covid-19-impfung.html; accessed on 28 June 2021), the Robert-Koch-Insitute (RKI) (https://impfdashboard.de/; accessed on 28 June 2021) and Github (https://github.com/ard-data/2020-rki-impf-archive/tree/master/data/9_csv_v2; accessed on 28 June 2021).

Acknowledgments: We would like to thank Samuel de Haas, Maximilian Gail, Peter Tillmann, Stefan Seifert, Walter Krämer and Jens Brunner for helpful comments. We thank Michael Kreil for continuously updating RKI-data on github.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

RKI	Robert-Koch-Institute
DEA	Data Envelopment Analysis
DMU	Decision making unit
CRS	Constant returns to scale
VRS	Variable returns to scale
BCC	Banker, Charnes and Cooper
DE	Germany
HB	Bremen
TH	Thuringia
BE	Berlin
BB	Brandenburg
SH	Schleswig-Holstein
NI	Lower Saxony
HH	Hamburg
BY	Bavaria
BW	Baden-Wuertemberg
HE	Hesse
RP	Rhineland Palatinate
ST	Saxony-Anhalt
NW	Northrhine-Westphalia
MV	Mecklenburg Western Pomerania
SN	Saxony
SL	Saarland

Appendix A. DEA

The influence of diminishing returns to scale was controlled for in the VRS DEA. However, there are further factors potentially affecting the results.

It is documented that there were occasional failures to deliver vaccines to certain areas (https://bit.ly/3uGEIul; accessed on 28 June 2021). Even though these are out of the control of the decision makers, it is unlikely that these delivery failures to individual federal states systematically affects the DEA scores over a period of over 4 months. The same holds for large-scale delivery failures of the vaccine of Astra-Zeneca (https://bit.ly/3jok2oe; accessed on 28 June 2021)). Moreover, the delivery failure of Astra-Zeneca's vaccine affected Germany as a whole so that it should not affect the relative efficiencies of the federal states' vaccination roll-outs.

One could argue that it is more difficult for the elderly to make and keep their appointments. This means that the broader the range of age cohorts eligible of being vaccinated the more strongly this could affect the DEA scores because, e.g., it requires more effort to vaccinate the elderly. In the first quarter the majority of vaccinations were given to the groups with highest (aged 80+) and second highest (aged 70–79) priority. Note that in the official regulation ("CoronaImpfV") priority groups within the population are not only defined by age but also by other factors, such as prior diseases, social relevance or exposition to infected people. The official regulation can be found here: https://bit.ly/2TbIpdR (accessed on 28 June 2021). It does not appear to be the case that demographics largely impact our results. Despite some overlaps, the federal states' age profiles (see https://bit.ly/349YNhb; accessed on 28 June 2021) do not seem to affect the results of our DEA. The following example illustrates this. If demographics had a significant impact on the relative success of the federal states' vaccination roll-outs, federal states with a younger population such as Hamburg (HH) with an average age of 42.1 would have an advantage over those with an older population such as Thuringia (TH) with an average age of 47.4 years. Hamburg and Thuringia have about the same population while Hamburg is significantly smaller in terms of area. However, according to the results presented in Tables 3 and 4 Hamburg is not unambiguously more efficient than Thuringia.

By the same token, our results do not seem to be largely driven by the federal states' geography. For instance, Bremen (HB), the most efficient federal state, and Saarland (SL), among the federal states with the lowest efficiency scores, are both relatively small federal states. However, the scores of the federal states with the lowest population density, Mecklenburg Western Pomerania (MV) and Brandenburg (BB) might be driven by the countries' geographies at least to some extent. Note that Mecklenburg Western Pomerania is assigned relatively high scores in model T (see Tables 1 and 2 in the main text), which means that the federal state performs relatively well despite its dispersed population, anyway. In these countries one would expect that it is relatively difficult for people in rural areas to reach vaccination centers and that doctor's offices are more efficient distributors of vaccination than large vaccination centers. Indeed, our analyses presented in Section 3.4 provide evidence that the integration of doctor's offices into those country's vaccination campaigns might have substantially increased efficiency. A fruitful extension would be to cluster the federal states. This would allow the researcher to identify the most suitable candidates for benchmarking.

A DEA with a different input is presented to demonstrate the robustness of the results presented in the main text. Tables A1 and A2 present the average DEA scores of CRS and VRS DEAs computed for the period 27 December 2020 to 16 May 2021 with only vaccine reserves in week $t-1$ as the input variable (rather than vaccine reserves in $t-1$ plus vaccine deliveries in week t). The results remain largely unchanged.

Table A1. Average efficiency scores for the period 27 December 2020 to 16 May 2021 for DEAs performed on a weekly basis under the CRS assumption using vaccine reserves at the end of the previous week as the input variable. Models T, 1S and 2S use total, first and second shots given as the respective output variable.

Federal State	T	1S	2S
BB	0.4286	0.3988	0.3399
BE	0.5822	0.562	0.6169
BW	0.4186	0.4276	0.3587
BY	0.609	0.6221	0.5104
HB	0.7961	0.7846	0.7055
HE	0.4081	0.4175	0.3945
HH	0.509	0.5238	0.4301
MV	0.6064	0.5613	0.4875
NI	0.3868	0.4055	0.3063
NW	0.4773	0.5112	0.3465
RP	0.6069	0.5752	0.4555
SH	0.5925	0.5711	0.43
SL	0.3999	0.4243	0.3289
SN	0.3788	0.3543	0.4116
ST	0.3934	0.3979	0.353
TH	0.5383	0.5257	0.5536

Table A2. Weekly average DEA efficiency scores for the period 27 December 2020 to 16 May 2021 under the VRS assumption using the vaccine stock at the end of the previous week as the input variable. Models T, 1S and 2S uses total, first and second shots given as the respective output variable.

Federal State	T	1S	2S
BB	0.5351	0.513	0.4278
BE	0.7466	0.7067	0.7372
BW	0.6372	0.6644	0.5538
BY	0.9306	0.902	0.7484
HB	0.9765	0.9765	0.9753
HE	0.5624	0.558	0.4916
HH	0.6115	0.6341	0.5095
MV	0.7152	0.7076	0.5812
NI	0.588	0.5977	0.4124
NW	1	0.947	0.8212
RP	0.7748	0.7257	0.6129
SH	0.7314	0.7187	0.5911
SL	0.4928	0.5245	0.4633
SN	0.4922	0.4505	0.5259
ST	0.4948	0.5052	0.4366
TH	0.643	0.6179	0.6849

Appendix B. Estimations

In the remainder of the Appendix, additional information to Section 3.4 are presented. First, we present an alternative estimation to analyze the effect of the integration of doctor's offices into Germany's vaccination campaign. The following fixed effects regression is estimated:

$$s_{i,t} = \alpha + \beta t + \gamma_i D_i \times t + \delta D_{\text{April 5}} + D_i + \epsilon_{i,t}. \tag{A1}$$

In Equation (A1), the share $s_{i,t}$ is explained by a time trend, a federal state dummy D_i interacted with the time trend and the dummy variable $D_{\text{April 5}}$ which takes the value 1 in the period starting on 5 April 2021 and 0 otherwise. Federal state-specific fixed effects are denoted D_i. With this specification, it is possible to disentangle how the share of vaccinations given in week t in relation to deliveries in week t and reserves in week $t-1$ evolves over time in the different federal states. The fixed effects are chosen such that the effect of the interaction term $D_i \times t$ has to be interpreted relative to the federal state Bremen (HB), as will be explained in more detail below. The results are presented in Table A3.

The results indicate that the share $s_{i,t}$ is significantly higher (1%-level) in the period after 5 April, which means that, on average, there is a structural break. The order of magnitude of the effect is the same as that presented in the main text. Note that this effect can be driven by doctor's offices being integrated into Germany's vaccination campaign, even though we cannot exclude that other factors also affected the results (see the discussion in the main text).

According to the results presented in Table A3 the share $s_{i,t}$ evolves over time on average. Every week, the share $s_{i,t}$ increases by 0.0108 plus the coefficient of the federal state dummy interacted with the time trend. This follows from $\frac{\partial s_{i,t}}{\partial t} = \beta + \gamma_i D_i$ based on Equation (A1). For instance, e.g., in Brandenburg (BB) the share of vaccinations given in week t in relation to deliveries in week t and reserves in week $t-1$ decreases by $0.0157 - 0.0108 = 0.0049$ per week (see Figure 3 in the main text).

Table A3. Output table for Equation (A1).

	Dep. Var. $s_{i,t}$	
$D_{\text{April 5}}$	0.116 **	(3.98)
Time trend	0.0108 ***	(5.86)
BB × Time trend	−0.0157 ***	(-3.17×10^{14})
BE × Time trend	−0.0152 ***	(-5.25×10^{14})
BW × Time trend	−0.00465 ***	(-1.62×10^{14})
BY × Time trend	−0.0138 ***	(-4.37×10^{14})
HE × Time trend	−0.00836 ***	(-1.58×10^{14})
HH × Time trend	−0.0111 ***	(-3.27×10^{14})
MV × Time trend	−0.0209 ***	(-4.46×10^{14})
NI × Time trend	−0.00235 ***	(-8.04×10^{13})
NW × Time trend	−0.000310 ***	(-1.08×10^{13})
RP × Time trend	−0.0168 ***	(-4.15×10^{14})
SH × Time trend	−0.0167 ***	(-5.78×10^{14})
SL × Time trend	−0.0105 ***	(-3.57×10^{14})
SN × Time trend	−0.00858 ***	(-2.98×10^{14})
ST × Time trend	−0.00943 ***	(-3.19×10^{14})
TH × Time trend	−0.00311 ***	(-1.08×10^{14})
Constant	0.333 ***	(26.80)
Observations	320	
R^2	0.424	
R^2 adjusted	0.360	

t statistics in parentheses; two and three asterisks correspond to $p < 0.01$ and $p < 0.01$, respectively.

To completely interpret the results, consider the federal state Bremen. The constant 0.333 can be interpreted as the starting value of the share $s_{\text{HB},t}$ for the federal state. Week $t = 1$ in our sample is the last week of December 2020. Subsequently, i.e., in calender weeks $t \in [1, 2, \ldots, 19]$ of 2021, that share increases by 0.0108 per week. For weeks 14–19, i.e., the calender weeks starting on 5 April 2021, the level of $s_{\text{HB},t}$ increases further by 0.116. Thus, in the last week, the average effect assigned to Bremen would be 0.6542.

Whereas we observe a clear time trend in the data for Bremen and, for instance, for Thuringia (TH) (see Figure 3 in the main text), other states do not exhibit such an obvious trend. In Equation (A1), we, therefore, still allow for a general time trend but do not impose such a trend for each federal state. In contrast, as a next step we approach the problem differently by allowing at the same time the analysis of whether the potential effect of the integration of doctor's offices into Germany's vaccination campaign differs between federal states.

$$s_{i,t} = \alpha + \beta t + \rho D_{\text{April 5}} + \eta_i D_i \times D_{\text{April 5}} + D_i + \epsilon_{i,t}. \tag{A2}$$

In Equation (A2), $s_{i,t}$ is explained by a time trend t, the variable $D_{\text{April 5}}$ that takes the value 1 for the period after 5 April 2021, and 0 otherwise as well as a federal state specific fixed effect D_i. The federal state dummy D_i is also interacted with $D_{\text{April 5}}$ to investigate whether a change in $s_{i,t}$ potentially triggered by the integration of doctor's offices into the vaccination campaign differs between federal states. This approach enables an investigation of whether the integration of doctor's offices into the vaccination campaign potentially improved efficiency of the vaccination campaign on average and whether this effects differs between federal states at the same time. The interaction term $D_i \times D_{\text{April 5}}$ is defined such that the results of each federal state have to be interpreted relative to Bremen. The results of the regression are presented in Table A4.

Table A4. Output table for Equation (A2).

	Dep. Var. $s_{i,t}$	
$D_{\text{April 5}}$	0.247 ***	(8.87)
$D_{\text{April 5}} \times$ BB	−0.206 ***	(-8.18×10^{13})
$D_{\text{April 5}} \times$ BE	−0.185 ***	(-7.26×10^{13})
$D_{\text{April 5}} \times$ BW	−0.0927 ***	(-3.74×10^{13})
$D_{\text{April 5}} \times$ BY	−0.181 ***	(-4.04×10^{13})
$D_{\text{April 5}} \times$ HE	−0.124 ***	(-4.77×10^{13})
$D_{\text{April 5}} \times$ HH	−0.158 ***	(-5.35×10^{13})
$D_{\text{April 5}} \times$ MV	−0.142 ***	(-5.20×10^{13})
$D_{\text{April 5}} \times$ NI	−0.0713 ***	(-2.87×10^{13})
$D_{\text{April 5}} \times$ NW	0.0120 ***	(4.82×10^{12})
$D_{\text{April 5}} \times$ RP	−0.186 ***	(-7.37×10^{13})
$D_{\text{April 5}} \times$ SH	−0.183 ***	(-6.61×10^{13})
$D_{\text{April 5}} \times$ SL	−0.198 ***	(-7.99×10^{13})
$D_{\text{April 5}} \times$ SN	−0.136 ***	(-5.44×10^{13})
$D_{\text{April 5}} \times$ ST	−0.108 ***	(-4.35×10^{13})
$D_{\text{April 5}} \times$ TH	−0.147 ***	(-5.93×10^{13})
Time trend	0.000941	(0.34)
Constant	0.333 ***	(14.03)
Observations	320	
R^2	0.40	
R^2 adjusted	0.332	

t statistics in parentheses; three asterisks correspond to $p < 0.01$.

One can see that the coefficient of the dummy $D_{\text{April 5}}$ is significant at the 0.1%-level. The coefficient of 0.247 reported in Table A4 implies that, all else equal, in the period after 5 April, 24.7% moreout of 100 available doses in a given week were vaccinated on average, or, in other words, 24.7% less doses were held back as reserves in Bremen. In interpreting the results, the potential effect of the integration of doctor's offices into the vaccination campaign is a jump in $s_{i,t}$ by 24.7% for Bremen. That jump is lower for all other federal states except for Northrhine-Westphalia, where the jump is $0.247 + 0.0120 = 0.259$. The lowest increase occurs in Brandenburg (BB) where the apparent effect of the integration of doctor's offices into the vaccination campaign is $0.247 - 0.206 = 0.041$.

A more detailed discussion on the differences between Regressions (A1) and (A2) is appropriate. The main difference is that (A1) allows the time trend to vary between federal states whereas (A2) allows the effect of the dummy $D_{\text{April 5}}$ to be federal state-specific. In contrast to (A1), in (A2) all differences in the time dimension between federal states is absorbed by the interaction term $D_i \times D_{\text{April 5}}$ so that the time trend eventually becomes insignificant. Based on a graphical analysis of Figure 3 in the main text one cannot clearly state which approach is more appropriate. For instance, Lower-Saxony (NI) and Hesse (HE) show similar patterns after 5 April, however, Lower-Saxony shows a clearer time trend. This indicates that Equation (A1) appears to be a more suitable approach to identify differences in the time dimension between federal states. On the other hand, comparing Bremen and Lower-Saxony, one can see that Bremen has a more pronounced effect of $D_{\text{April 5}}$. One could argue that there is a more visible structural break in Bremen than in Lower-Saxony. In that case, Equation (A2) seems more appropriate to identify differences in the time dimension. Based on these observations, it seems appropriate to discuss both variations. In any case, both estimations indicate structural breaks after 5 April, so that the results presented here confirm the findings outlined in the main text.

Finally, it can be analyzed which federal states show a time trend and for whom a structural break after 5 April can be diagnosed. In doing so, we estimate the following regression for each federal state:

$$s_{i,t} = \alpha_i + \beta_i t + \rho_i D_{\text{April 5}} + \gamma_i D_{\text{April 5}} \times t. \quad \text{(A3)}$$

In Equation (A3), we test whether each federal state has a time trend, a structural break after April 5 and whether the time trend changes after April 5. The results are presented in Table A5.

Table A5. Time trend and structural break after 5 April 2021, and the interaction between the time trend and the structural break in $s_{i,t}$ for each federal state. The number of stars indicates the p-values of the respective t-statistics. In contrast to the previous analyses, *** indicates the $p \leq 0.01$, ** $p \leq 0.05$ and * $p \leq 0.1$.

Federal State	Time Trend (t)	Structural Break ($D_{\text{April 5}}$)	Interaction ($D_{\text{April 5}} \times t$)
BB	none	positive **	none
BE	none	positive ***	negative ***
BW	none	positive **	none
BY	none	positive *	none
HB	none	positive ***	negative **
HE	none	none	none
HH	none	positive *	none
MV	negative ***	none	positive **
NI	none	none	none
NW	none	positive ***	none
RP	none	none	none
SH	none	none	none
SL	none	none	none
SN	none	positive ***	negative ***
ST	none	positive ***	negative **
TH	positive **	positive ***	negative ***

Table A5 shows that 10 out of 16 federal states exhibit a structural break after 5 April 2021. That structural break is positive, i.e., an upward jump in $s_{i,t}$ is observed. However, starting from that higher level, one can see that in 5 of these federal states (HB, BE, SN, ST, TH) the shares $s_{i,t}$ start to decline again, which follows from the negative sign of the coefficient γ_i of the interaction term $D_{\text{April 5}} \times t$. This indicates that in these federal states there was a strong initial impetus on $s_{i,t}$ at the beginning of 5 April, which began to level out afterwards. In the remaining 5 federal states (BB, BW, BY, HH, NW) the shares remain at a higher level. For Mecklenburg Western Pomerania (MV), we do not find a structural break, however, a negative time trend is found with an opposing positive trend after 5 April (positive coefficient of the interaction term). No effect of the integration of doctor's offices into the vaccination campaign is found for 5 federal states (HE, NI, RP, SH, SL).

As a further robustness check, a Zivot-Andrews structural break test [38] was performed. This can be seen as a conservative approach because the test identifies endogenously at most one structural break. Based on the data, a structural break was identified in calender weeks 13 and 14 (i.e., the weeks around 5 April) in 6 federal states (including, in particular, Bremen and Northrhine-Westphalia) despite the scarcity of observations. These findings are consistent with the interpretation outlined in the main text that the identified effect is driven by those federal states where the integration of general practitioners into the vaccination campaign had a particularly strong effect.

References

1. WHO. Novel Coronavirus (2019-nCoV). Situation Report—22, Data as Reported by 11 February 2020. Technical Report, World Health Organization. 2020. Available online: https://apps.who.int/iris/handle/10665/330991 (accessed on 28 June 2021).
2. Hua, J.; Shaw, R. Corona virus (Covid-19)"infodemic" and emerging issues through a data lens: The case of china. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2309. [CrossRef]
3. El Zowalaty, M.E.; Järhult, J.D. From SARS to COVID-19: A previously unknown SARS-related coronavirus (SARS-CoV-2) of pandemic potential infecting humans–Call for a One Health approach. *ONE Health* **2020**, *9*, 100124. [CrossRef]
4. Mallapaty, S. Can COVID vaccines stop transmission? Scientists race to find answers. *Nat. News* **2021**. [CrossRef]
5. Bubar, K.M.; Reinholt, K.; Kissler, S.M.; Lipsitch, M.; Cobey, S.; Grad, Y.H.; Larremore, D.B. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **2021**, *371*, 916–921. [CrossRef] [PubMed]
6. Jahn, B.; Sroczynski, G.; Bicher, M.; Rippinger, C.; Mühlberger, N.; Santamaria, J.; Urach, C.; Schomaker, M.; Stojkov, I.; Schmid, D.; et al. Targeted COVID-19 Vaccination (TAV-COVID) Considering Limited Vaccination Capacities—An Agent-Based Modeling Evaluation. *Vaccines* **2021**, *9*, 434. [CrossRef] [PubMed]
7. McKee, M.; Rajan, S. What can we learn from Israel's rapid roll out of COVID 19 vaccination? *Isr. J. Health Policy Res.* **2021**, *10*, 5. [CrossRef] [PubMed]
8. Mills, M.C.; Salisbury, D. The challenges of distributing COVID-19 vaccinations. *EClinicalMedicine* **2021**, *31*, 10067. [CrossRef]
9. Goel, R.K.; Nelson, M.A. Drivers of COVID-19 vaccinations: Vaccine delivery and delivery efficiency in the United States. *NETNOMICS Econ. Res. Electron. Netw.* **2021**. [CrossRef]
10. ECDC. Rollout of COVID-19 Vaccines in the EU/EEA: Challenges and Good Practice. Technical Report, European Centre for Disease Prevention and Control. 2021. Available online: https://www.ecdc.europa.eu/en/publications-data/rollout-covid-19-vaccines-eueea-challenges-and-good-practice (accessed on 28 June 2021).
11. Rossman, H.; Shilo, S.; Meir, T.; Gorfine, M.; Shalit, U.; Segal, E. COVID-19 dynamics after a national immunization program in Israel. *Nat. Med.* **2021**, *27*, 1055–1061. [CrossRef] [PubMed]
12. Vilches, T.N.; Zhang, K.; Van Exan, R.; Langley, J.M.; Moghadas, S.M. Projecting the impact of a two-dose COVID-19 vaccination campaign in Ontario, Canada. *Vaccine* **2021**, *39*, 2360–2365. [CrossRef] [PubMed]
13. Pageaud, S.; Ponthus, N.; Gauchon, R.; Pothier, C.; Rigotti, C.; Eyraud-Loisel, A.; Bertoglio, J.P.; Bienvenüe, A.; Gueyffier, F.; Vanhems, P.; et al. Adapting French COVID-19 vaccination campaign duration to variant dissemination. *medRxiv* **2021**. [CrossRef]
14. Ratzan, S.; Schneider, E.C.; Hatch, H.; Cacchione, J. Missing the Point—How Primary Care Can Overcome Covid-19 Vaccine "Hesitancy". *N. Engl. J. Med.* **2021**, *384*, e100. [CrossRef]
15. Charnes, A.; Cooper, W.W.; Rhodes, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **1978**, *2*, 429–444. [CrossRef]
16. Thanassoulis, E. DEA and its use in the regulation of water companies. *Eur. J. Oper. Res.* **2000**, *127*, 1–13. [CrossRef]
17. Agrell, P.J.; Bogetoft, P.; Tind, J. DEA and dynamic yardstick competition in Scandinavian electricity distribution. *J. Product. Anal.* **2005**, *23*, 173–201. [CrossRef]
18. Shi, F.X.; Lim, S.H.; Chi, J. Railroad productivity analysis: Case of the American Class I railroads. *Int. J. Product. Perform. Manag.* **2011**, *60*, 372–386. [CrossRef]
19. Staat, M. Efficiency of hospitals in Germany: A DEA-bootstrap approach. *Appl. Econ.* **2006**, *38*, 2255–2263. [CrossRef]
20. Kounetas, K.; Papathanassopoulos, F. How efficient are Greek hospitals? A case study using a double bootstrap DEA approach. *Eur. J. Health Econ.* **2013**, *14*, 979–994. [CrossRef] [PubMed]
21. Valdmanis, V.; Walker, D.; Fox-Rushby, J. Are vaccination sites in Bangladesh scale efficient? *Int. J. Technol. Assess. Health Care* **2003**, *19*, 692–697. [CrossRef]
22. Birch, S.; Gafni, A. Cost effectiveness/utility analyses: Do current decision rules lead us to where we want to be? *J. Health Econ.* **1992**, *11*, 279–296. [CrossRef]
23. Nas, T.F. *Cost-Benefit Analysis: Theory and Application*; Lexington Books: Lanham, MD, USA, 2016.
24. Aldamak, A.; Zolfaghari, S. Review of efficiency ranking methods in data envelopment analysis. *Measurement* **2017**, *106*, 161–172. [CrossRef]
25. Cooper, W.W.; Seiford, L.M.; Zhu, J. *Handbook on Data Envelopment Analysis*; Springer: Berlin, Germany, 2011.
26. Banker, R.D.; Charnes, A.; Cooper, W.W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manag. Sci.* **1984**, *30*, 1078–1092. [CrossRef]
27. Torres, I.; Lopez-Cevallos, D.; Artaza, O.; Profeta, B.; Kang, J.; Machado, C.V. Vaccine scarcity in LMICs is a failure of global solidarity and multilateral instruments. *Lancet* **2021**, *397*, 1804. [CrossRef]

28. Rosen, B.; Waitzberg, R.; Israeli, A. Israel's rapid rollout of vaccinations for COVID-19. *Isr. J. Health Policy Res.* **2021**, *10*, 6. [CrossRef]
29. Shilo, S.; Rossman, H.; Segal, E. Signals of hope: Gauging the impact of a rapid national vaccination campaign. *Nat. Rev. Immunol.* **2021**, *21*, 198–199. [CrossRef] [PubMed]
30. Cylus, J.; Panteli, D.; van Ginneken, E. Who should be vaccinated first? Comparing vaccine prioritization strategies in Israel and European countries using the Covid-19 Health System Response Monitor. *Isr. J. Health Policy Res.* **2021**, *10*, 1–3. [CrossRef] [PubMed]
31. Tewarson, H.; Greene, K.; Fraser, M.R. State Strategies for Addressing Barriers during the Early US COVID-19 Vaccination Campaign. *Am. J. Public Health* **2021**, *111*, 1073–1077. [CrossRef] [PubMed]
32. Goel, R.K.; Nelson, M.A. COVID-19 internet vaccination information and vaccine administration: Evidence from the United States. *J. Econ. Financ.* **2021**. [CrossRef]
33. Sim, F. Early Covid-19 vaccination rollout: A commentary from England. *Isr. J. Health Policy Res.* **2021**, *10*, 1–4. [CrossRef]
34. Jentzsch, A.; Bleckwenn, M. Corona vaccination in a family practice: A first experience report. *MMW Fortschritte Med.* **2021**, *163*, 12–17. [CrossRef]
35. Caplin, A.; Leahy, J. Economic Theory and the World of Practice: A Celebration of the (S, s) Model. *J. Econ. Perspect.* **2010**, *24*, 183–202. [CrossRef]
36. Geoffroy, F.; Traulsen, A.; Uecker, H. Vaccination strategies when vaccines are scarce: On conflicts between reducing the burden and avoiding the evolution of escape mutants. *medRxiv* **2021**. [CrossRef]
37. Voysey, M.; Clemens, S.A.C.; Madhi, S.A.; Weckx, L.Y.; Folegatti, P.M.; Aley, P.K.; Angus, B.; Baillie, V.L.; Barnabas, S.L.; Bhorat, Q.E.; et al. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: A pooled analysis of four randomised trials. *Lancet* **2021**, *397*, 881–891. [CrossRef]
38. Zivot, E.; Andrews, D.W.K. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *J. Bus. Econ. Stat.* **2002**, *20*, 25–44. [CrossRef]

6 Health Communication and COVID-19 Vaccine Hesitancy: A Synthetic Control Approach

Author: Phil-Adrian Klotz

Presented at the following conferences:

- 9. International Conference on Risk Analysis, Perugia (Italy)
- 2022 French Stata Conference in Applied Econometrics, Marseille (France)
- 2022 VfS Annual Conference, Basel (Switzerland)

Health Communication and COVID-19 Vaccine Hesitancy: A Synthetic Control Approach

Phil-Adrian Klotz*

January 25, 2024

Abstract

This paper empirically analyzes the impact of health communication on the COVID-19 vaccination rate using a quasi-experimental design. Based on a speech by the President of France, Emmanuel Macron, we examine how political leaders can influence the willingness of their citizens to get vaccinated by transmitting scientific insights into a clear and vivid message as well as by threatening credibly with future restrictions for unvaccinated people. By applying the synthetic control method it is shown that a televised address by Macron has increased the vaccination rate in France by roughly 7.5%. We test the robustness of this result by applying an event study design as well as the synthetic Difference-in-Differences approach. Our findings imply that health communication is an effective weapon to change the beliefs of unvaccinated citizens and to overcome COVID-19 vaccine hesitancy.

Keywords: COVID-19, health policy, vaccination, synthetic control

JEL Classification: H42, I18, Z18

*Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen, Germany. E-mail: phil.a.klotz@wirtschaft.uni-giessen.de. Tel.: +49-641 99 22060. I would like to thank participants at the Stata Conference in Applied Econometrics 2022, the VfS Annual Conference 2022, and specifically John Antonakis, Maximilian Maurice Gail, Georg Götz and Daniel Herold for their very helpful comments.

Declarations

Funding: No funding was received to assist with the preparation of this manuscript.

Conflict of interest: The author has no competing interests to declare that are relevant to the content of this article.

Ethics approval: Not applicable.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and material: The data set generated during and/or analysed during the current study is available from the corresponding author on reasonable request.

Code availability: All code for data cleaning and analysis is available on request.

1 Introduction

"A summer of mobilization for vaccination. That is what we must aim for." These words, which were spoken in a televised address by the President of France, Emmanuel Macron, were heard live by more than 22 million French citizens in July, 2021.[1] Thereby, the French President tried to exert his influence by motivating people to get vaccinated against COVID-19 in a time of increasing vaccine hesitancy. Beside emphasizing the importance of getting vaccinated, Macron also announced mandatory vaccination for healthcare workers and future restrictions for unvaccinated people in his speech.

Our work is an empirical investigation that seeks to answer the question whether the televised address by Macron has had an impact on the willingness of French citizens to get vaccinated against COVID-19. Hence, we examine the impact of leadership communication on follower behavior in an economically relevant situation. This is an interesting research area because even though leadership research has shown that leaders cannot only influence follower behavior by implementing financial incentives but also through personal abilities to persuade and motivate (House, 1977), causal evidence on this topic is scarce and mixed so far (Antonakis et al., 2021).

Most of the empirical research on leader communication has been done in the field of applied psychology using questionnaires to gauge perceptions of charisma (Alimo-Metcalfe & Alban-Metcalfe, 2001; Bass, 1985; Podsakoff et al., 1990). However, those studies cannot identify causal effects since questionnaire measures of charisma are endogenous (biased by omitted variables)

[1] https://www.france24.com/en/europe/20210712-follow-live-france-s-macron-addresses-the-nation-as-covid-19-delta-variant-surges.

as well as conceptually and empirically confounded (Antonakis et al., 2010, 2016; Fischer et al., 2020). In contrast, our empirical study can provide causal inference on leadership communication since we use a quasi-experimental design and do not rely on survey data.

The COVID-19 crisis is characterized by a high degree of volatility, uncertainty, complexity, and ambiguity so that leadership has become very important. Political leaders should exert their influence by promoting what science has to say, clearly communicate evidence-based policies and inducing individuals to act in ways that benefit the collective (Antonakis, 2021; Chou & Budenz, 2020; Galasso et al., 2022; Rzymski et al., 2021; Soares et al., 2021).

Different studies have figured out that leader's charisma - the ability to transmit information in a symbolic, value-based, and emotional manner - can get individuals to undertake personally costly but socially beneficial actions to solve coordination problems in the end (Antonakis et al., 2021; Bastardoz & Van Vugt, 2019). With respect to the COVID-19 crisis, Jensen et al. (2021) have shown that charisma of US governor speeches had significantly increased physical distancing of US citizens. Hence, beside altering the cost-benefit calculation of individuals about whether to receive the vaccine, in times of a relatively high vaccine hesitancy it is also a key challenge for policy makers to change the beliefs of unvaccinated citizens by using communication solutions (Bates et al., 2022; Zhou et al., 2022).

Another important tool for political leaders, which might accelerate the vaccination campaign, is the credible threat of future restrictions for unvaccinated people (Walkowiak et al., 2021). Such strict, punitive measures could be a restricted access to restaurants, sport facilities, indoor events or public transport for people without a proof of vaccination or recovery. This is a form of coercive power, whose strength crucially depends on the magnitude of the negative valence of the threatened punishment as well as the opportunity whether the citizens can avoid the punishment by conformity (French et al., 1959). To be effective, it is also important that political leaders have a high credibility and that there is a sufficient public trust in the government measures (Bargain & Aminjonov, 2020; Gesser-Edelsburg et al., 2020; Liu & Huang, 2021; Tabler et al., 2023). When leaders lack credibility, followers can easily ignore their announcements (Wilson & Rhodes, 1997).

However, the empirical evidence on the effectiveness of using sanctions or the threat of sanctions is limited and mixed so far. There is only some evidence from public goods experiments showing that the implementation of a (costly) sanctioning system, which punishes free-riding persons, can reduce the number of free-riders and raise the willingness to cooperate (altruistic punishment, see Balafoutas et al. (2014), Fehr and Gächter (2000, 2002), and Yamagishi (1986)). Most of the empirical evidence from applied psychology would suggest that punishment (or its threat) actually reduces the performance of people (Podsakoff et al., 1982, 2006).

The results in this paper suggest that health communication in combination with the threat of sanctions can positively affect people's behavior. We examine the impact of Macron's speech on the daily vaccination rate in France by using the synthetic control method and find a positive and significant effect. Thereby, we rely on 13 other European countries as a reference group in our donor pool. We conclude that leaders cannot only affect follower behavior through the design of incentives and institutions, but also through charismatic leadership and coercive power.

The rest of the paper is structured as follows. In Section 2, we give some background on the COVID-19 crisis and present descriptive statistics on the vaccination campaigns in France and the 13 control countries. We introduce the synthetic control method in Section 3.1 and outline the empirical results in Section 3.2. In Section 4, we conclude with the policy implications of our research.

2 Vaccination Campaigns

The administrative and logistical challenges of distributing the COVID-19 vaccinations have been substantial in most of the European countries. The majority of the countries in our data set have a relatively old population and a health care system which is not as efficient as for example the digitized health care system of Israel (Rosen et al., 2021). In Section 2.1, we provide some background on the COVID-19 crisis as well as on the vaccination campaigns in general. Then, we present some descriptive statistics on the vaccination campaigns in France and the other 13 European countries in Section 2.2.

2.1 Background

Almost four years after the World Health Organization was informed about the outbreak of the SARS-Cov-2 virus in Wuhan (China), the COVID-19 crisis still has an impact on people's health and the economy worldwide. By the time this paper was written, there have been roughly 772 million confirmed cases and almost 7 million fatalities in connection with COVID-19 throughout the world.[2]

Vaccination campaigns have started around the world in the end of 2020. Vaccination is a central pillar to bring the pandemic under control if a sufficient number of people are vaccinated.[3] The four primary COVID-19 vaccines in Europe have been shown to be effective at preventing COVID-19 infections, hospitalizations, and deaths (Jabłońska et al., 2021; Liang et al., 2021).[4] As of November 29, 2023, approx. 75 percent of the people in the European Union have received at least one vaccine dose, with large variation across the 27 countries.[5]

Although the used vaccines are free of charge and highly effective, vaccine hesitancy remains prevalent in many states (Dror et al., 2020; Lazarus et al., 2021; Sallam, 2021). Despite ample evidence that vaccination reduces the probability of infection, hospitalization and death, there are many people who choose not to get vaccinated (Thorp, 2020). Possible reasons are manifold. For instance, concerns about the vaccine safety, effectiveness and side effects contribute to vaccine hesitancy (Courbage & Peter, 2021; Fisher et al., 2020; Koskan et al., 2023; Vásquez et al., 2021). In a poll conducted in the United States in September 2020, 62% of the respondents worry that political pressure would lead to a rush to approve vaccines without proper

[2] See https://covid19.who.int/ (data as of November 29, 2023).

[3] Models predict that at least 70 percent of the population have to be vaccinated in order for the incidence of infection to decline (Randolph & Barreiro, 2020).

[4] For instances, the Robert Koch Institute (RKI) has stated that vaccinations during the "third wave" prevented over 706,000 new cases of infection and 38,300 deaths in Germany (see https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2021/Ausgaben/35_21.pdf?__blob=publicationFile).

[5] See https://ourworldindata.org/covid-vaccinations.

attention to safety and effectiveness (Newsroom, 2020). Beyond, many individuals share false information about the vaccines (e.g., about their alleged side effects) on social media (Marco-Franco et al., 2021).

From an economics perspective, vaccination against COVID-19 poses a positive externality: an individual who decides to get vaccinated does not only increase its own utility due to a decreased probability of getting infected, but also confers significant benefits (spillover benefits) on others. For instance, vaccination reduces the virus transmission (Shah et al., 2021) and leads to a lower probability of a hospitalization which is associated with a lower burden for the health care system (Nasreen et al., 2021).[6] Against this background, economic theory predicts that vaccination remains at a suboptimal level because positive externalities are not internalized (Chen & Toxvaerd, 2014). This justifies governmental intervention (Stiglitz, 1988). The national governments organize the vaccine procurement and allocation as well as subsidize the vaccinations (most of the developed countries offer vaccination free of charge for their citizens). Moreover, some policy makers have implemented incentive programs to nudge unvaccinated citizens to receive a COVID-19 vaccine (e.g., the "Vax-a-Million" lottery in Ohio, see Brehm et al. (2021)).

The COVID-19 vaccination roll-out in France has started on December

[6] In the later course of the pandemic, the Omicron variant has become dominant in almost all countries. Even though the protection after the primary two dose series of the approved vaccines is lower for this variant compared to previous COVID-19 variants (Pajon et al., 2022), it has been shown that three exposures to the spike protein of SARS-CoV-2 by either infection or vaccination can also induce high-quality antibodies against Omicron (Wratil et al., 2022). In the meantime, also updated vaccines against more recent Omicron subvariants have been released (see https://www.statnews.com/2022/11/04/pfizer-biontech-report-bivalent-covid-19-vaccine-more-protective-than-original-vaccine/). So, vaccination is still the most effective way to protect people against an infection.

27, 2020. As in most of the countries of the EU, mainly the four vaccines of BioNTech-Pfizer, Moderna, AstraZeneca, and Johnson&Johnson have been used. France as well as other European countries have implemented prioritization schemes at the beginning of the vaccination campaigns so that especially elderly and more vulnerable people as well as essential workers were vaccinated first.[7] Especially mobile vaccination teams have been deployed to inject the elderly residents at the early stage of the vaccination campaigns before many countries started to open large vaccination centers in spring 2021.[8] One advantage of the French campaign was the early integration of general practitioners compared to, for example, Germany. By the end of February 2021, approximately 30,000 general practitioners in France were authorized to administer vaccines, whereas their German counterparts were integrated only in April 2021.

2.2 Data and Descriptive Statistics

We use a panel data set recording data on daily vaccinations at the country level from "Our World in Data", publicly available on https://github.com/owid/COVID-19-data/tree/master/public/data (Mathieu et al., 2021). Besides data on vaccinations, this data set also includes data on COVID-19 cases, deaths, hospitalizations and testing as well as socio-economic variables. Data on COVID-19 vaccinations is available for 218 countries.[9] Our period

[7]These governmental decisions on a prioritization scheme were in line with the views in representative surveys at that time (See, e.g., Luyten et al., 2022).

[8]See https://www.connexionfrance.com/article/French-news/Where-are-the-Covid-19-vaccination-centres-in-France and https://www.tagesschau.de/thema/impfzentren/.

[9]However, there exists significant variation in data quality among individual countries, with noticeable gaps in vaccination data for several countries. Hence, we only incorporate 14 European countries in our empirical analysis characterized by high-quality data (see

of investigation starts on December 27, 2020 (when the vaccination campaign has begun in France) and ends on September 9, 2021.

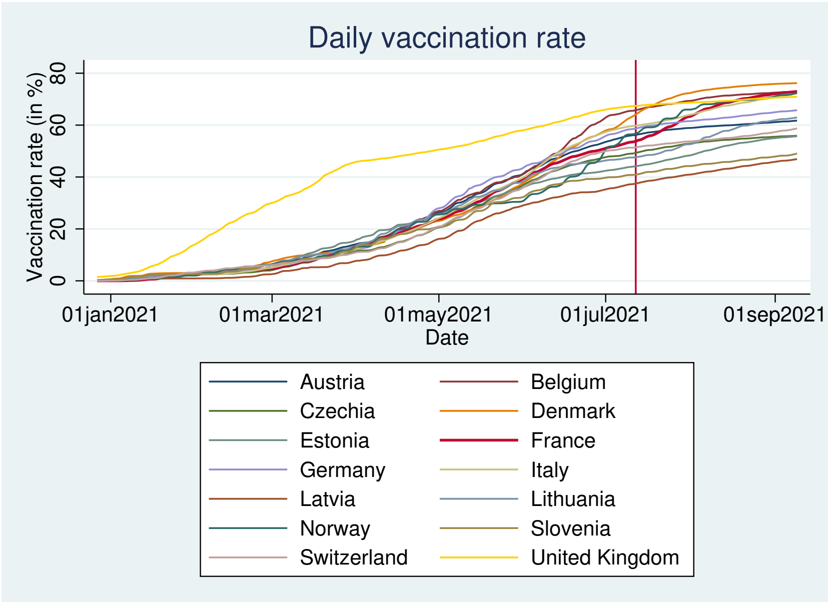

Figure 1: Progress of vaccination rates in France and the 13 control countries (vertical line represents televised address of Macron on July 12, 2021, Source: Our World in Data).

The development of the share of people vaccinated at least once in our observation period is plotted in Figure 1 for the 14 countries in our data set.[10] It is obvious that the vaccination campaigns have been very slow at the beginning of 2021 in most of the countries, which was mainly driven by two reasons. First, there was a vaccine shortage due to the failed joint-procurement approach of the EU member states leading to a situation in which the EU member countries had relatively less vaccine doses compared to, e.g., the United Kingdom or the US.[11] Second, the campaigns were slowed

Section 3).

[10]In the meantime, many people have also received a second vaccine dose as well as booster and second booster vaccinations. However, since the research topic of this paper is vaccine hesitancy, we are only focused on the first vaccine shot.

[11]See https://www.economist.com/briefing/2021/03/31/why-the-eus-covid-19-vaccination-programme-went-wrong.

down relatively early due to the suspension of the AstraZeneca vaccine in many countries.[12] Moreover, also a high degree of bureaucracy and relatively few trust of residents in their health institutions have been issues.[13] All those factors led to a situation in which only 13% of the French and only 12% of the German population got at least one dose of the COVID-19 vaccine by the end of March, while, e.g., in the UK already 46% of the people had a first shot at that time.

In the following month, the vaccination speed has increased more rapidly in the European countries (see Figure 1), especially due to the risen vaccine supply. However, vaccine hesitancy has become an issue in many parts of Europe in the summer of 2021. While in the early stage of the campaign vaccine shortage was a problem, now the demand for vaccinations was lower than the supply and the vaccine reserves increased drastically week by week. For instance, 16.27 million doses were in the German vaccine stock in calendar week 30 (July 26 – August 1, 2021).[14] This constitutes jabs for approx. 20% of Germany's population.

The declining vaccination rates in these months are also clearly visible in Figure 1. The curves of the individual countries started to grow concavely at the beginning of the summer 2021. Yet, about the time of the televised address by Emmanuel Macron (red vertical line in Figure 1) there is a visible

[12]See https://www.economist.com/briefing/2021/03/31/why-the-eus-covid-19-vaccination-programme-went-wrong and https://www.rki.de/DE/Content/Kommissionen/STIKO/Empfehlungen/AstraZeneca-Impfstoff-2021-03-30.html. For instance, Deiana et al. (2022) find empirical evidence that the suspension of the AstraZeneca vaccine has slowed down the vaccination campaign in Italy.

[13]See Götz et al. (2021) and https://www.economist.com/europe/2021/01/30/why-frances-vaccination-roll-out-has-been-so-slow.

[14]See https://impfdashboard.de/en/.

rise in the daily vaccination rate of France (red curve) so that the share of vaccinated people has become larger in France compared to most of the other countries in our data set during the summer months.

On July 12, 2021, President Emmanuel Macron gave a televised address broadcasted by all important French television channels and watched by more than 22 million citizens. In his speech, he particularly emphasized the importance of getting vaccinated ("A summer of mobilization for vaccination. That is what we must aim for.") and announced the introduction of a COVID-19 passport. From August 2021 on, only those who were fully vaccinated, recovered from COVID-19 or had a negative test result would be allowed into cinemas, sport stadiums, restaurants, bars, shopping centers, nightclubs or on long-distance trains and flights. Beyond, Macron stated that vaccination will be mandatory for all health workers (enforced from September on) and that COVID-19 tests will no longer be free from October 15, 2021 on.[15]

In the hours following Mr. Macron's announcement, over 1 million citizens booked vaccination appointments via the widely-used medical app Doctolib. Three-fifths of the bookings have been for people aged between 18 and 39 years.[16] On July 28, 2021, France overtook Germany in terms of people who received at least one shot and later this figure has also been higher than, e.g., in the US or the UK. On the day of Mr. Macron's announcement, 54% of the French citizens had received at least one dose, by December 15, 2021 this figure was around 78%.

[15] https://www.economist.com/europe/2021/09/18/how-france-tackled-vaccine-hesitancy.

[16] https://www.economist.com/graphic-detail/2021/07/14/why-vaccine-shy-french-are-suddenly-rushing-to-get-jabbed.

	Mean	S.D.	Min.	Max.
A: Data on Covid-19 vaccination				
Share of people vaccinated (in %)	29.78	23.54	0.0005	76.18
Share of people fully vaccinated (in %)	20.3	20.2	0.000005	73.67
B: Time-varying predictors				
New Covid-19 cases 7 days ago	218.9	236.31	0	1745.04
New Covid-19 cases 14 days ago	219.36	238.38	0	1745.04
Stringency index	58.88	15.39	23.15	87.96
C: National demographic variables				
Median age	43.32	2.06	39.7	47.9
Persons aged 70 and above (in %)	13.29	1.42	10.81	16.24
Life expectancy	80.7	2.46	75.29	83.78
GDP per capita	40276.6	10767.37	25063.85	64800.06
Population density (inhabitants/km^2)	145.22	101.2	14.46	375.56

Table 1: Summary statistics.

Table 1 shows summary statistics for variables we employ for our main analysis. The variables are measured at the country level. Panel A contains the variables related to measuring the Covid-19 vaccination progress. Panel B shows information on the time-varing predictors, new Covid-19 cases and the strigency index, and panel C displays all predictors related to the country's demographic structure.

3 Empirical Analysis

The goal of our empirical analysis is to analyze the impact of the Macron speech on the vaccination rate in France. In order to construct a reliable counterfactual scenario, we employ the synthetic control method. This design is presented in Section 3.1. The results of our empirical approach are then outlined in Section 3.2.

3.1 Synthetic Control Method

In order to examine whether the televised address by Emmanuel Macron from July 12, 2021 was effective in increasing the willingness of French citizens to get vaccinated against COVID-19, we apply the synthetic control method (SCM). This method was proposed for the causal assessment of policy interventions based on aggregate outcome measures (Abadie & Gardeazabal, 2003; Abadie et al., 2010).[17] Synthetic control models optimally choose a set of weights for the untreated units to produce an optimally estimated counterfactual to the unit that received the treatment. This "synthetic control" then displays what would have happened to the treated unit if the treatment had never occurred. It is a powerful generalization of the DiD approach and some authors have argued that it was one of the most important contributions to quantitative comparative case studies (Athey & Imbens, 2017). Compared to the canonical DiD approach, this method also has the advantages of not relying on the parallel trend assumption and that the control group is not selected *ad hoc* and subjectively.

The synthetic control method is based on the assumption that a combination of control groups often does a better job than using a single comparison unit alone. Hence, we construct a synthetic France as the convex combination of 13 European countries in our donor pool. The synthetic France in this method is selected to be the weighted average of all comparison units that best resemble the characteristics of real France in the pre-treatment

[17] It has also be proven to be a useful method to study the impacts of the Covid-19 pandemic, e.g., to analyze the effect of making face mask mandatory or to quantify the effect of lockdown measures (Born et al., 2021; Mitze et al., 2020).

period. To construct the counterfactual trajectory of vaccination rates for France in the absence of the Macron speech, we use six pre-speech average vaccination rates, one pre-speech average share of fully vaccinated people, two lagged variables for the number of infected people, the median age, the share of people aged over 70, life expectancy, GDP per capita, the population density, and a COVID-19 stringency index as covariates.

Following the notation of Cunningham (2021), the matching variables X_1 and X_0 are chosen as predictors of post-intervention outcomes and must be unaffected by the intervention. The weights are chosen so as to minimize the norm $||X_1 - X_0W||$, given the two constraints that (I) no unit receives a negative weight and (II) that the sum of all weights must equal one. Defining V as some $(k \times k)$ symmetric and positive semidefinite matrix and X_{jm} as the value of the mth covariates for unit j, the synthetic control weights minimize:

$$\sum_{m=1}^{k} v_m \left(X_{1m} - \sum_{j=2}^{J+1} w_j X_{jm} \right)^2, \tag{1}$$

where v_m is a weight that reflects the relative importance that we assign to the mth variable when we calculate the difference between France and the synthetic control.

The choice of V is important since W^* depends on its choice. The synthetic control $W^*(V)$ reproduces the behavior of the outcome variable for France in the absence of the Macron speech. Hence, the weights $v_1, ..., v_k$ should reflect the predictive value of the covariates. In this paper, V is chosen by minimizing the mean squared prediction error:

$$\sum_{t=1}^{T_0}\left(Y_{1t}-\sum_{j=2}^{J+1}w_j^*(V)Y_{jt}\right)^2, \tag{2}$$

where Y_{jt} is the vaccination rate of unit j at time t. Once we find the weights solving our minimization problem, we can multiply them by the control units at all time periods to get a synthetic control for France:

$$y_{sc} = Y_{co}\hat{w}^{sc}. \tag{3}$$

Then, the average treatment effect on the treated (ATT) is simply the average of the value for France minus the average of the synthetic control in the post-treatment period:

$$\hat{\tau} = \bar{y}_{post,fr} - \bar{y}_{post,sc}. \tag{4}$$

3.2 Results

The results from our synthetic control method indicate that the Macron speech had an significant effect on the vaccination rate in France. In Figure 2, the vaccination rates of France and the 13 European countries in our donor pool are plotted. Apart from the United Kingdom, the vaccination rates of the observed countries start from similar levels in our observation period and also have a very similar development in the pre-treatment period. Based on the 13 control states, the synthetic control method produces the optimally estimated synthetic France.

The outcome on the construction of the synthetic France is displayed

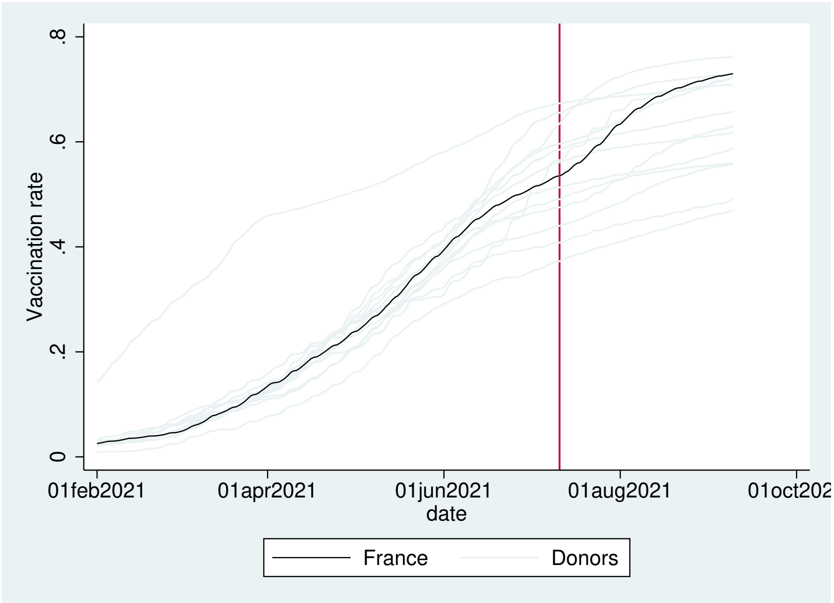

Figure 2: Vaccination rates in France and the 13 countries in our donor pool.

in Table 2, which compares the pre-treatment characteristics of the actual France with that of the synthetic France as well as with the simple average of the 13 countries in the donor pool. Obviously, the simple average of the 13 control countries does not seem to provide an appropriate control group for France. In particular, the measures mapping the pandemic situation ($Cases_{t-1}$, $Cases_{t-2}$, stringency index) as well as the vaccination rates prior to the Macron speech were very different in the 13 control states compared to France. In contrast, the synthetic France seems to produce an accurate control group based on the used covariates.

Variables	France Real	Synthetic	Average of 13 control states
$Cases_{t-1}$	285.3792	294.7625	189.6514
$Cases_{t-2}$	297.1168	309.6421	193.2084
Median Age	42	43.2088	43.42308
Percent aged 70+	13.079	12.75462	13.30908
Life Expectancy	82.66	80.57409	80.54385
GDP per capita	38605.67	39181.31	40405.13
Population density	122.578	150.2196	146.9605
Stringency index	61.94	61.75605	56.79161
Percent fully vaccinated June 29	30.5	31.62	33.71
Percent vaccinated February 1	2.54	2.46	3.34
Percent vaccinated March 1	4.71	4.99	7.34
Percent vaccinated April 1	13.49	13.17	15.11
Percent vaccinated May 1	23.84	24.1	25.73
Percent vaccinated June 1	39.41	38.36	38.62
Percent vaccinated July 5	51.945	52.07	52.06

Table 2: Vaccination rate predictor means.

The weights of each control country in the synthetic France calculated based on equations (1) and (2) are presented in Table 3. The calculated weights indicate that the vaccination rate in France is best reproduced by a combination of eight different European countries. Five states in the donor pool are assigned zero weights.

State	Weight
Austria	0.238
Belgium	0.158
Czechia	0.331
Denmark	0.005
Estonia	0.157
Germany	0
Italy	0.033
Latvia	0
Lithuania	0
Norway	0.045
Slovenia	0
Switzerland	0.033
United Kingdom	0

Table 3: Country weights in the synthetic France.

The country weights constituting our synthetic control are also presented

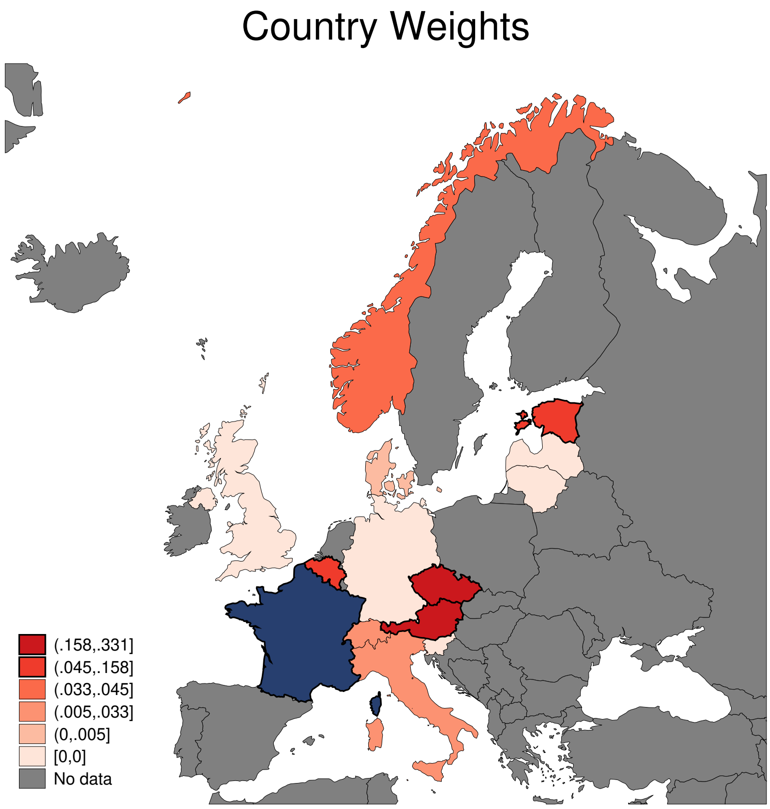

Figure 3: Country Weights for synthetic France.

in Figure 3. Most of the control countries receiving high weights are geographically very close to France. The four countries that receive the largest weights (Czechia, Austria, Belgium and Estonia) are also members of the European Union so that they have comparable socio-economic structures. These countries are also very similar with respect to the Covid-19 vaccination roll-out as they used mostly the same types of vaccines and were also affected by the failed joint-procurement approach of the EU member states (see Section 2).

In Figure 4, we plot the daily vaccination rates (people who got at least one vaccine dose) for France and its synthetic counterpart during the period

February 1 to September 9, 2021.[18] Notice that the synthetic France very closely tracks the trend of the French vaccination rate in the pre-speech period, suggesting that the synthetic France is an appropriate approximation for the real vaccination rate in France. Shortly after the Macron speech on July 12 (red vertical line), the two lines in Figure 4 begin to diverge noticeably.

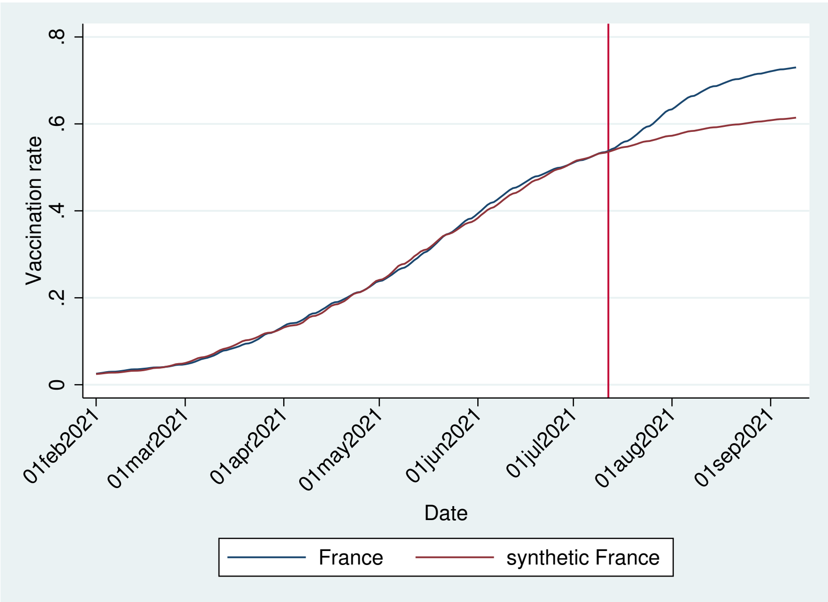

Figure 4: Trends in vaccination rates: France vs. synthetic France.

This obvious divergence between the vaccination rates of the real and the synthetic France is even better illustrated in Figure 5. Here, the daily gap in the vaccination rate between France and its synthetic counterpart is plotted. While these gaps are very closely to zero in the pre-speech period, the increase in the gap after July 12 suggests that the Macron speech had a relatively large effect on the French vaccination rate. Based on equation (4),

[18]We shorten the observation period for the synthetic control method since for some countries in our donor pool there is no vaccination data available before February 1, 2021.

we can calculate the ATT of our synthetic control method which is equal to 7.4%.

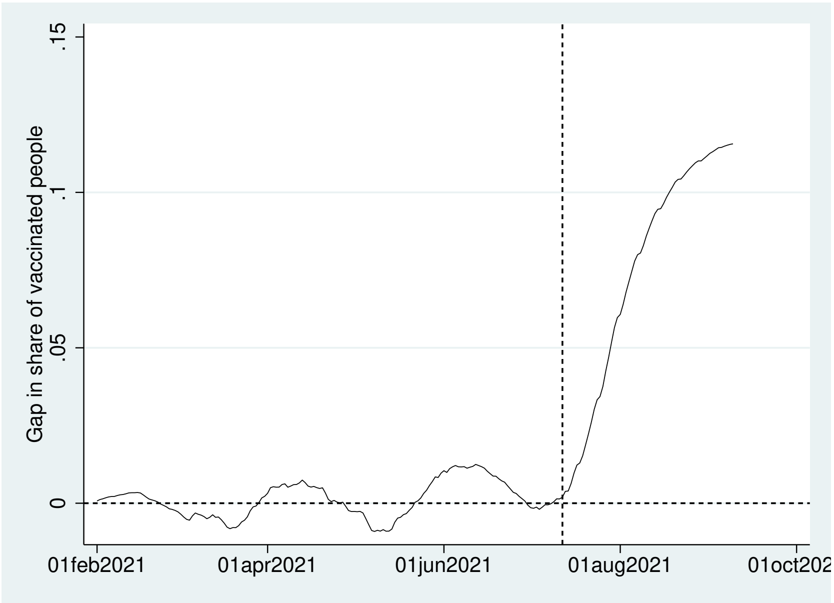

Figure 5: Gap in the vaccination rate between France and synthetic France.

Finally, we evaluate the gap in the vaccination rate between France and the 13 control countries by considering the distributions of the ratios of post-/pre-speech root mean squared prediction errors (RMSPE). Therefore, in a first step we calculate the RMSPE for the pre- and post-speech period for France as well as for the 13 countries in our donor pool based on the following equation:

$$RMSPE = \left(\frac{1}{T - T_0} \sum_{t=T_0+t}^{T} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{\frac{1}{2}}, \tag{5}$$

where Y_{jt} is the vaccination rate of unit j at time t and w_j^* is a vector of optimally chosen weights. Then, we compute the ratios of post- to pre-speech RMSPE to get a standard of comparison. Figure 6 displays the distribution of those post-/pre-speech ratios of the RMSPE for France and the 13 control countries. The ratio for France clearly stands out in the figure as no control country achieves such a large ratio. The post-speech RMSPE for France is about 14 times the RMSPE of the pre-speech.

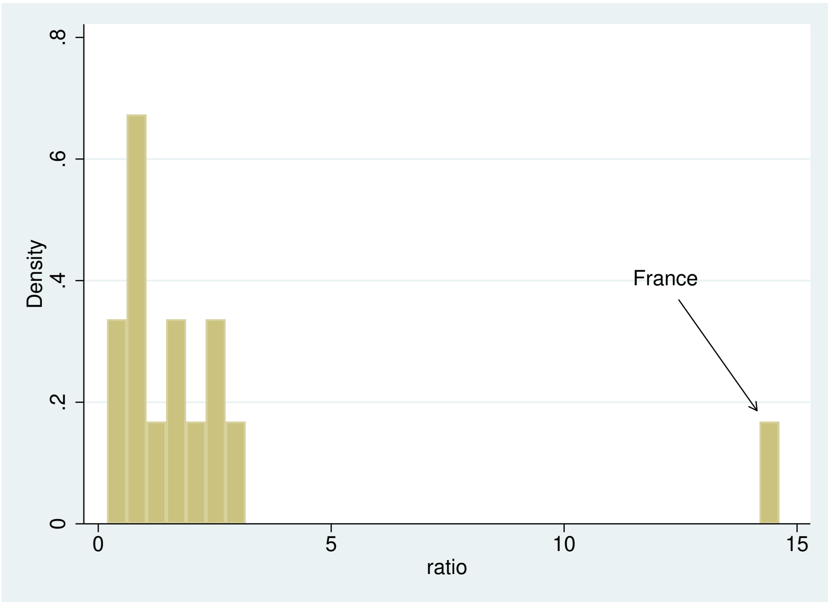

Figure 6: Ratio of post-speech RMSPE and pre-speech RMSPE: France and the 13 control countries.

To check the robustness of our main results, we evaluate the importance of individual countries and predictors to construct the synthetic twin of France. Hence, we construct six alternative synthetic twins for France by respectively dropping the four countries from our donor pool with the largest weights in our baseline approach (Czechia, Austria, Belgium, Estonia) and by using two adjusted sets of predictors (one without using the vaccination rates in different points of time in the pre-treatment period as predictors and one without the number of lagged Covid-19 cases).

Figure 7 shows the pre-treatment fit and the post-treatment effect of the Macron speech under these adjusted synthetic twins for France. As the

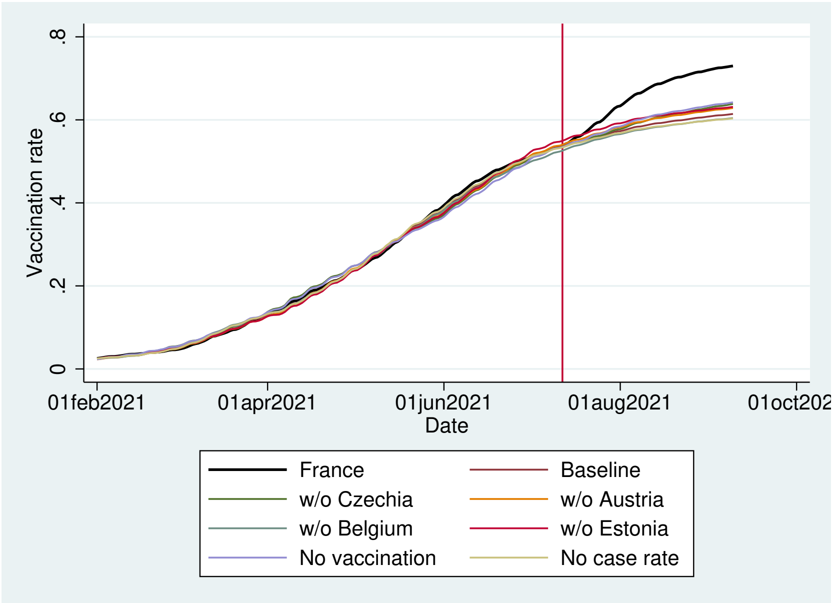

Figure 7: Alternative contructions of the synthetic France.

figure indicates, excluding any countries from the donor pool and changing the predictor sets does not affect our main findings.

We also inquired into the robustness of our results by employing an event study approach as well as the synthetic Difference-in-Differences design. The results are qualitatively identical to our SCM findings and are presented in Appendix A and B.

4 Conclusion

This article asked whether political leaders can influence the decision of their citizens to get vaccinated against COVID-19. Since vaccine hesitancy is omnipresent in many countries, it is a key challenge for policy makers to nudge a reluctant population. Beside designing incentive programs to encourage COVID-19 vaccinations, changing the beliefs of unvaccinated citizens by transmitting scientific insights is an important task for political leaders. They should counter false information on social media by providing secure knowledge on the safety and effectiveness of the approved vaccines. Moreover, the credible threat of future restrictions for unvaccinated residents might further increase the people's willingness to get vaccinated.

The results of our synthetic control approach imply that a televised speech by France's President Emmanuel Macron has increased the willingness of French citizens to get vaccinated by roughly 7.5%. President Macron has pointed out the importance to get vaccinated and announced extensive future restrictions for unvaccinated residents in his well-received speech. Several placebo tests, an event study design as well as a synthetic Difference-in-Differences approach also prove that the speech has significantly increased the vaccination rate in France.

Our findings are consistent with works exploring the impact of leadership communication on solving coordination problems (see, e.g., Antonakis et al. (2021)) and some newer studies that point out the effect of leader's charisma on the compliance with the COVID-19 rules (see, e.g., Jensen et al. (2021)). However, it remains a task for future research to find the optimal remedies to raise the vaccination rate in different countries (Bates et al., 2022). Possibly, in countries outside France leadership communication is less effective compared to, e.g., financial incentives.

Of course, a limitation of our study is that the motivation effect and the effect of the threat on future restrictions are confounded. In our approach we cannot identify whether the transmission of scientific insights in Macron's speech or the announcement of future restrictions for unvaccinated people has significantly increased the willingness to get vaccinated in the French population (or only the combination of both). So future research in this area should try to estimate the effect of both leadership "styles" independently of each other.

In some countries like Germany, there has been a debate on the introduction of a statewide general mandatory vaccination against COVID-19. Since there are good reasons to believe that the willingness to get vaccinated might not be sufficiently high to avoid non-pharmaceutical interventions (NPIs) like lockdowns in the future, some policy makers have thought that this is the last remedy to reach the last pockets of unvaccinated people. However, our results suggest that applying health communication can also be an effective weapon to change the beliefs of unvaccinated citizens and can possibly avoid the necessity of a general mandatory vaccination.

References

Abadie A., Diamond A., & Hainmueller J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie A., & Gardeazabal J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Alimo-Metcalfe B., & Alban-Metcalfe R. J. (2001). The development of a new transformational leadership questionnaire. *Journal of occupational and organizational psychology*, *74*(1), 1–27.

Antonakis J. (2021). Leadership to defeat covid-19. *Group Processes & Intergroup Relations*, *24*(2), 210–215.

Antonakis J., Bastardoz N., Jacquart P., & Shamir B. (2016). Charisma: An ill-defined and ill-measured gift.

Antonakis J., Bendahan S., Jacquart P., & Lalive R. (2010). On making causal claims: A review and recommendations. *The leadership quarterly*, *21*(6), 1086–1120.

Antonakis J., d'Adda G., Weber R. A., & Zehnder C. (2021). "just words? just speeches?" on the economic value of charismatic leadership. *Management Science*.

Arkhangelsky D., Athey S., Hirshberg D. A., Imbens G. W., & Wager S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118.

Athey S., & Imbens G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic perspectives*, *31*(2), 3–32.

Balafoutas L., Nikiforakis N., & Rockenbach B. (2014). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences*, *111*(45), 15924–15927.

Bargain O., & Aminjonov U. (2020). Trust and compliance to public health policies in times of covid-19. *Journal of public economics*, *192*, 104316.

Bass B. M. (1985). *Leadership and performance beyond expectations*. Collier Macmillan.

Bastardoz N., & Van Vugt M. (2019). The nature of followership: Evolutionary analysis and review. *The Leadership Quarterly*, *30*(1), 81–95.

Bates B. R., Villegas-Botero A., Costales J. A., Moncayo A. L., Tami A., Carvajal A., & Grijalva M. J. (2022). Covid-19 vaccine hesitancy in three latin american countries: Reasons given for not becoming vaccinated in colombia, ecuador, and venezuela. *Health Communication*, *37*(12), 1465–1475.

Born B., Dietrich A. M., & Müller G. J. (2021). The lockdown effect: A counterfactual for sweden. *Plos one*, *16*(4), e0249732.

Brehm M. E., Brehm P. A., & Saavedra M. (2021). *The ohio vaccine lottery and starting vaccination rates* (tech. rep.). Mimeo.

Chen F., & Toxvaerd F. (2014). The economics of vaccination. *Journal of theoretical biology*, *363*, 105–117.

Chou W.-Y. S., & Budenz A. (2020). Considering emotion in covid-19 vaccine communication: Addressing vaccine hesitancy and fostering vaccine confidence. *Health communication*, *35*(14), 1718–1722.

Courbage C., & Peter R. (2021). On the effect of uncertainty on personal vaccination decisions. *Health Economics*, *30*(11), 2937–2942.

Cunningham S. (2021). *Causal inference*. Yale University Press.

Deiana C., Geraci A., Mazzarella G., & Sabatini F. (2022). Perceived risk and vaccine hesitancy: Quasi-experimental evidence from italy. *Health Economics*, *31*(6), 1266–1275.

Diederichs M., van Ewijk R., Isphording I. E., & Pestel N. (2022). Schools under mandatory testing can mitigate the spread of sars-cov-2. *Proceedings of the National Academy of Sciences*, *119*(26), e2201724119.

Dror A. A., Eisenbach N., Taiber S., Morozov N. G., Mizrachi M., Zigron A., Srouji S., & Sela E. (2020). Vaccine hesitancy: The next challenge in the fight against covid-19. *European journal of epidemiology*, *35*(8), 775–779.

Fehr E., & Gächter S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980–994.

Fehr E., & Gächter S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fischer T., Hambrick D. C., Sajons G. B., & Van Quaquebeke N. (2020). Beyond the ritualized use of questionnaires: Toward a science of actual behaviors and psychological states. *The Leadership Quarterly*, *31*(4), 101449.

Fisher K. A., Bloomstone S. J., Walder J., Crawford S., Fouayzi H., & Mazor K. M. (2020). Attitudes toward a potential sars-cov-2 vaccine: A survey of us adults. *Annals of internal medicine*, *173*(12), 964–973.

French J. R., Raven B., & Cartwright D. (1959). The bases of social power. *Classics of organization theory*, *7*, 311–320.

Galasso V., Pons V., Profeta P., Becher M., Brouard S., & Foucault M. (2022). From anti-vax intentions to vaccination: Panel and experimental evidence from nine countries.

Gesser-Edelsburg A., Cohen R., Hijazi R., Shahbari N. A. E., et al. (2020). Analysis of public perception of the israeli government's early emergency instructions regarding covid-19: Online survey study. *Journal of medical Internet research*, *22*(5), e19370.

Götz G., Herold D., Klotz P.-A., & Schäfer J. T. (2021). Efficiency in covid-19 vaccination campaigns—a comparison across germany's federal states. *Vaccines*, *9*(7).

House R. J. (1977). A theory of charismatic leadership. *Leadership: The cutting egde*.

Jabłońska K., Aballéa S., & Toumi M. (2021). The real-life impact of vaccination on covid-19 mortality in europe and israel. *medRxiv*.

Jensen U., Rohner D., Bornet O., Carron D., Garner P., Loupi D., & Antonakis J. (2021). Combating covid-19 with charisma: Evidence on governor speeches and physical distancing in the united states.

Koskan A. M., Teeter B. S., Daniel C. L., LoCoco I. E., Jensen U. T., & Ayers S. L. (2023). Us adults' reasons for changing their degree of willingness to vaccinate against covid-19. *Journal of Public Health*, 1–13.

Lazarus J. V., Ratzan S. C., Palayew A., Gostin L. O., Larson H. J., Rabin K., Kimball S., & El-Mohandes A. (2021). A global survey of potential acceptance of a covid-19 vaccine. *Nature medicine*, *27*(2), 225–228.

Liang L.-L., Kuo H.-S., Ho H. J., & Wu C.-Y. (2021). Covid-19 vaccinations are associated with reduced fatality rates: Evidence from cross-county quasi-experiments. *Journal of Global Health*, *11*.

Liu W., & Huang Y. (2021). Does relationship matter during a health crisis: Examining the role of local government-public relationship in the public acceptance of covid-19 vaccines. *Health Communication*, 1–11.

Luyten J., Tubeuf S., & Kessels R. (2022). Rationing of a scarce life-saving resource: Public preferences for prioritizing covid-19 vaccination. *Health Economics*, *31*(2), 342–362.

Maneenop S., & Kotcharin S. (2020). The impacts of covid-19 on the global airline industry: An event study approach. *Journal of air transport management*, *89*, 101920.

Marco-Franco J. E., Pita-Barros P., Vivas-Orts D., González-de-Julián S., & Vivas-Consuelo D. (2021). Covid-19, fake news, and vaccines: Should regulation be implemented? *International Journal of Environmental Research and Public Health*, *18*(2), 744.

Mathieu E., Ritchie H., Ortiz-Ospina E., Roser M., Hasell J., Appel C., Giattino C., & Rodés-Guirao L. (2021). A global database of covid-19 vaccinations. *Nature human behaviour*, *5*(7), 947–953.

Mitze T., Kosfeld R., Rode J., & Wälde K. (2020). Face masks considerably reduce covid-19 cases in germany. *Proceedings of the National Academy of Sciences*, *117*(51), 32293–32301.

Nasreen S., He S., Chung H., Brown K. A., Gubbay J. B., Buchan S. A., Wilson S. E., Sundaram M. E., Fell D. B., Chen B., et al. (2021). Effectiveness of covid-19 vaccines against variants of concern, canada. *Medrxiv*.

Newsroom K. (2020). Poll: Most americans worry political pressure will lead to premature approval of a covid-19 vaccine. half say they would not get a free vaccine approved before election day. *KFF; Kaiser Family Foundation: San Francisco, CA, USA*.

Pajon R., Doria-Rose N. A., Shen X., Schmidt S. D., O'Dell S., McDanal C., Feng W., Tong J., Eaton A., Maglinao M., et al. (2022). Sars-cov-2 omicron variant neutralization after mrna-1273 booster vaccination. *New England Journal of Medicine*.

Podsakoff P. M., Bommer W. H., Podsakoff N. P., & MacKenzie S. B. (2006). Relationships between leader reward and punishment behavior and subordinate attitudes, perceptions, and behaviors: A meta-analytic review of existing and new research. *Organizational Behavior and Human Decision Processes*, *99*(2), 113–142.

Podsakoff P. M., MacKenzie S. B., Moorman R. H., & Fetter R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *The leadership quarterly*, *1*(2), 107–142.

Podsakoff P. M., Todor W. M., & Skov R. (1982). Effects of leader contingent and noncontingent reward and punishment behaviors on subordinate performance and satisfaction. *Academy of management journal*, *25*(4), 810–821.

Randolph H. E., & Barreiro L. B. (2020). Herd immunity: Understanding covid-19. *Immunity*, *52*(5), 737–741.

Rosen B., Waitzberg R., & Israeli A. (2021). Israel's rapid rollout of vaccinations for covid-19. *Israel journal of health policy research*, *10*(1), 1–14.

Rzymski P., Borkowski L., Drag M., Flisiak R., Jemielity J., Krajewski J., Mastalerz-Migas A., Matyja A., Pyrć K., Simon K., et al. (2021). The strategies to support the covid-19 vaccination with evidence-based communication and tackling misinformation. *Vaccines*, *9*(2), 109.

Sallam M. (2021). Covid-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates. *Vaccines*, *9*(2), 160.

Shah A. S., Gribben C., Bishop J., Hanlon P., Caldwell D., Wood R., Reid M., McMenamin J., Goldberg D., Stockton D., et al. (2021). Effect of vaccination on transmission of sars-cov-2. *New England Journal of Medicine*, *385*(18), 1718–1720.

Soares P., Rocha J. V., Moniz M., Gama A., Laires P. A., Pedro A. R., Dias S., Leite A., & Nunes C. (2021). Factors associated with covid-19 vaccine hesitancy. *Vaccines*, *9*(3), 300.

Stiglitz J. E. (1988). Economics of the public sector new york. *NY: WW Norton & Company.*

Tabler J., Snyder J. A., White C., Freng A., & Thunström L. (2023). Covid-19 health practices and attitudes in the united states: The role of trust in healthcare. *Journal of Public Health*, *31*(10), 1631–1644.

Thorp H. H. (2020). Persuasive words are not enough. *Science*, *368*(6498), 1405–1405. https://doi.org/10.1126/science.abd4085

Vásquez W. F., Trudeau J. M., & Alicea-Planas J. (2021). Immediate and informative feedback during a pandemic: Using stated preference analysis to predict vaccine uptake rates. *Health Economics*, *30*(12), 3123–3137.

Walkowiak M. P., Walkowiak J. B., & Walkowiak D. (2021). Covid-19 passport as a factor determining the success of national vaccination campaigns: Does it work? the case of lithuania vs. poland. *Vaccines*, *9*(12), 1498.

Wilson R. K., & Rhodes C. M. (1997). Leadership and credibility in n-person coordination games. *Journal of Conflict Resolution*, *41*(6), 767–791.

Wratil P. R., Stern M., Priller A., Willmann A., Almanzar G., Vogel E., Feuerherd M., Cheng C.-C., Yazici S., Christa C., et al. (2022). Three exposures to the spike protein of sars-cov-2 by either infection or vaccination elicit superior neutralizing immunity to all variants of concern. *Nature Medicine*, 1–1.

Yamagishi T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and social Psychology*, *51*(1), 110.

Zhou S., Villalobos J. P., Munoz A., & Bull S. (2022). Ethnic minorities' perceptions of covid-19 vaccines and challenges in the pandemic: A qualitative study to inform covid-19 prevention interventions. *Health Communication*, *37*(12), 1476–1487.

A Event Study

This section describes the results of an event study design, which is our first robustness check for the main SCM results. The event study design (dynamic Difference-in-Differences model) is a frequently applied tool to evaluate policy treatments (see, e.g., Cunningham (2021), ch. 9.4) and also has already been used to identify COVID-19 policies (see, e.g., Diederichs et al. (2022) and Maneenop and Kotcharin (2020)). One advantage of this approach is a higher flexibility compared to the SCM estimation in terms of controlling for heterogeneous time trends during the pre- and also during the treatment period.

The key identification assumption of our DiD approach is that, in the absence of the Macron speech, the trends in the daily vaccination rates would have been the same in France and the 13 control countries. The event study allows for a pattern in vaccination rates leading up and following the speech of Macron so that we can see whether there is an effect, and how long it lasts. We include a set of interaction terms comprised of an indicator for whether the country is France and indicators for each of the 40 days before and 40 days after the Macron speech. The week before the televised speech is the omitted group, yielding the following equation:

$$
\begin{aligned}
v_{it} = \gamma_i + \lambda_t + \delta Cases_{i,t-1} + FR_i \times [\sum_{k=-40}^{-2} \pi_k 1(Day_t = k) \\
+ \sum_{k=0}^{40} \rho_k 1(Day_t = k)] + \epsilon_{it},
\end{aligned}
\tag{6}
$$

where v_{it} is the share of people who received at least one vaccine dose in country i and day t. The variable λ_t depicts the day fixed effects and γ_i represents country-level fixed effects accounting for country-level characteristics that do not vary over time. $Cases_{i,t-1}$ is a covariate which depicts the lagged number of new COVID-19 cases (per million inhabitants, lagged by one week) to control for different pandemic events in the observed countries. The coefficients π_k and ρ_k provide the estimated change in vaccination rates relative to the week before Macron's speech ($k = -1$). Estimates close to zero on the interaction term π_k in the pre-speech weeks provide evidence against concerning pre-trends.

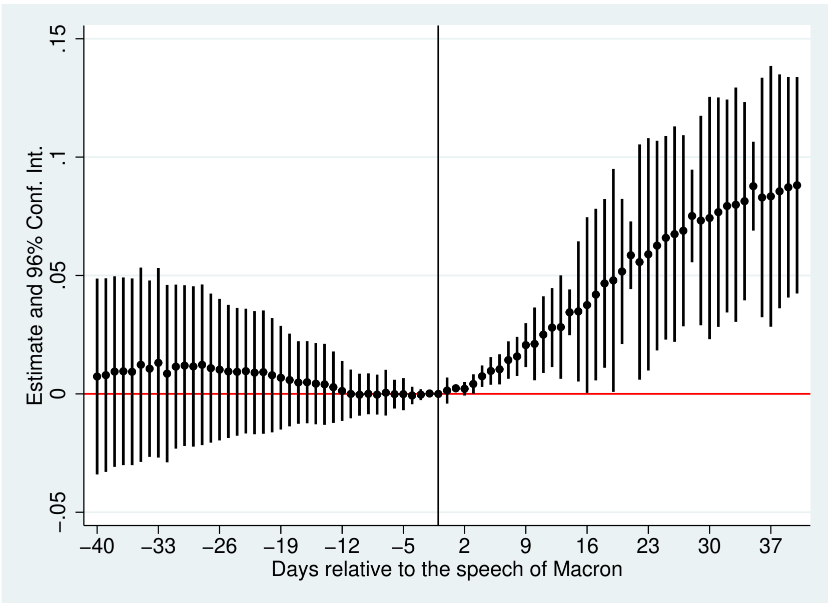

Figure 8: Event Study results based on Equation (6). The dependent variable is the share of people who received at least on vaccine dose (using 95% confidence intervals). The vertical black line indicates the day before Macron's speech using daily bins.

Figure 8 presents the estimates from the event study design in equation (6). The estimated coefficients π_k and ρ_k (see equation (6)) and their 95% confidence intervals are plotted. The estimated effects are relative to the day before Macron's speech and suggest that the speech has significantly increased the daily vaccination rate in France after the speech. The coefficients for the first few days after the speech might be insignificant because even if some residents have made a vaccination appointment just after the televised speech, it would have taken some time to get the jab.

Moreover, the pre-treatment coefficients π_k are close to the zero line and statistically insignificant, providing empirical support for the necessary identification assumption of parallel trends in the absence of the Macron speech. Afterwards, there is a clear increase starting approx. four days after the Macron speech, and the post-treatment coefficients ρ_k become significant at day $t = 3$. Taken together, our robustness check based on an event study design qualitatively support our main SCM results.

B Synthetic Difference-in-Differences Design

In another robustness check, we apply the synthetic Difference-in-Differences (SDID) approach as introduced by Arkhangelsky et al. (2021). This method combines aspects of the DiD and the SCM and, thus, is invariant to additive unit-level shifts and allows for valid large-panel inference (like DiD) as well as reweights and combines pre-treatment trends to weaken the reliance on the parallel trend assumption.

We still have a balanced panel with $N = 14$ countries and $T = 221$ days. Again, we denote v_{it} our outcome for country i in period t and our binary treatment is now given by $W_{it} \in \{0, 1\}$. Relating to SCM, $\hat{\omega}^{sdid}$ gives the weights which align the pre-treatment trends in the outcome of the untreated countries with those for France. We use these weights in a basic two-way fixed effects regression to estimate the average causal effect of the Macron speech (denoted by τ):

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \Big\{ \sum_{i=1}^{N=14} \sum_{t=1}^{T=220} (v_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \Big\}. \tag{7}$$

In equation (7), $\hat{\lambda}_t^{sdid}$ denotes the time weights which balance the pre-treatment with the post-treatment time periods are also included. By using country and time weights for the SDID estimator $\hat{\tau}^{sdid}$, countries and time periods which are on average similar to France, respectively its periods, are weighted more strongly. The inclusion of using only similar units and similar periods makes the estimator more robust relative to a standard DiD estimator.

Figure 9 illustrates the result of this synthetic DiD approach. SDID reweights the unexposed control countries to make their time trend close to France pre-treatment and then applies a DID analysis to this reweighted panel. In the period before the speech of Macron, the vaccination rate of the control group is very similar to the one of France. However, the vaccination rate in France significantly rises over the level of the control countries following the televised speech of Macron. The ATT of our synthetic DiD approach is at 7%, which is very close to our SCM main result.

Finally, Figure 10 plots the differences in the vaccination rate outcome from the synthetic DiD approach for each control country, where the size

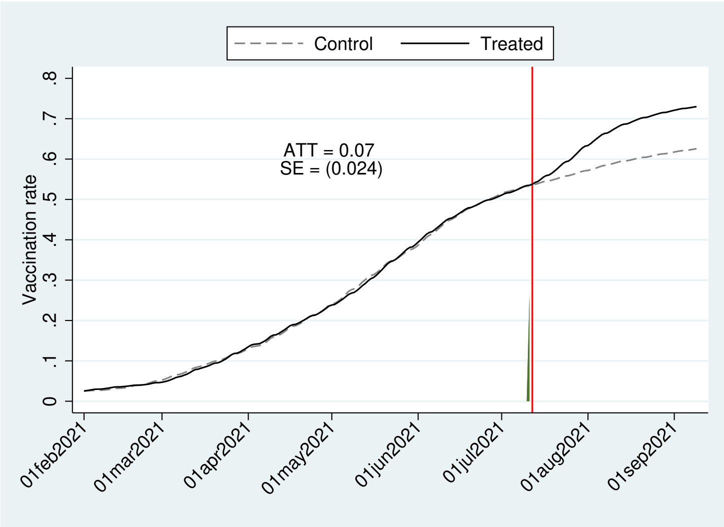

Figure 9: Trends in the vaccination rate over time for France and the relevant weighted average of control states, with the weights used to average pre-treatment time periods at the bottom of the graphs.

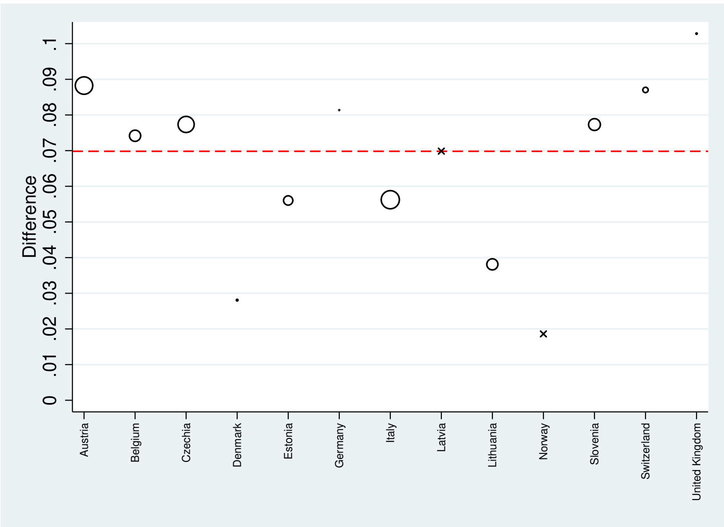

Figure 10: State-by-state adjusted outcome difference in the vaccination rate from the synthetic DiD approach. The weights $\hat{\omega}^{sdid}$ are indicated by dot size. The weighted average of these differences are illustrated by the horizontal line. States are ordered alphabetically. Observations with zero weight are denoted by an $\times$ symbol.

of each point corresponds to its weight $\hat{\omega}_i^{sdid}$ and countries with zero weight are denoted by an $\times$ symbol. The red dashed horizontal line illustrates the estimated average effect of these differences. Obviously, SDID does not give any control country particularly high influence, suggesting that after weighting, we have achieved the desired parallel trends without inducing excessive variance in the estimator.

7 Conclusion

This doctoral thesis is structured in two main parts: In the first main part, two papers address competition restraints in the book market. The third paper of this first part analyzes the pass-through of temporary fuel tax reductions in the European gasoline market. The second main part of this thesis focuses on the German and European vaccination campaigns during the COVID-19 crisis. In all five papers of this thesis, causal inference methods are employed to identify and measure causal links.

In the social sciences, randomized experiments are oftentimes not feasible because one cannot randomly assign individuals into treatment and control groups. Hence, researchers in this field frequently rely on natural experiments to estimate treatment effects, where assignment to the treatment is not randomized by a researcher but decided by a policy maker. Thereby, the basic Difference-in-Differences (DiD) design has become the most-commonly used technique in economics to estimate treatment effects (De Chaisemartin & d'Haultfoeuille, 2022). However, most recent research in this area has indicated that the standard DiD estimator is biased if the treatment effect is not constant across locations and over time (Goodman-Bacon, 2021). Hence, newer causal inference methods have been explored which provide unbiased estimator under heterogeneous treatment effects.

In the five papers of this doctoral thesis, well-established causal inference methods like the basic DiD design or the instrumental variable approach as well as newer methods like the staggered DiD approach or the synthetic control method are employed to answer policy relevant research questions in the fields of competition policy and health economics.

In the first main part of this thesis, the results of the two papers on the book market imply that competition restraints may not automatically result in anti-competitive effects. The first paper *The Substitutability Between Brick-and-Mortar Stores and e-Commerce - The Case of Books* implies that a larger number of physical bookstores promotes book sales, while in this market the e-Commerce seems to be no perfect substitute to physical retailers from the viewpoint of the consumers. Thus, fixed book price systems like in the German book market may support the policy goals of securing a broad supply of books. The results of the second paper *The Impact of the Agency Model on E-book Prices: Evidence from the UK* suggest that vertical market restraints can have pro-competitive effects under certain conditions. In this paper, the effect of using agency arrangements between suppliers and retailers on consumer prices in the British e-Book market is analyzed. Based on an unique dataset containing bestselling and long tail e-Books, it is shown that digital book prices are significantly lower under the agency model compared to the traditional wholesale model.

Also the third paper of this thesis focuses on a competition policy topic. In the paper *Pass-through of temporary fuel tax reductions: Evidence from Europe*, the pass-through rates of temporary fuel tax reductions are estimated and the effect on retail margins of the petroleum companies is analyzed. In general, the results in this paper imply a full-shifting of the 2022 temporary fuel tax reductions in the three largest European economies, even though there are heterogeneous effects over time. In line with this, these fuel tax reductions had no significant effects on the retail margins of the petroleum companies. From a competition policy perspective, the estimated pass-through rates imply that the alleged competition restrictions in the European gasoline market can at least not hinder a high pass-through of temporary tax reductions.

The second part of this doctoral thesis contains two papers in the field of empirical health economies. Both papers empirically analyze how different policies can increase the speed of COVID-19 vaccination campaigns. In the paper *Efficiency in COVID-19 Vaccination Campaigns - A Comparison across Germany's Federal States*, the efficiency of the initial vaccination campaigns in the 16 German federals states is systematically evaluated. The results in this paper are two-fold: first, the paper identifies remarkable differences in the speed of the vaccination campaigns between the individual federal states. Federal state-specific policies like the collaboration between private firms and public officials have caused these differences. Second, the results in this paper imply that the integration of doctors' offices into the campaign has sped up the vaccination roll-out in Germany.

The paper *Health Communication and COVID-19 Vaccine Hesitancy: A Synthetic Control Approach* examines whether science-based health communication of political leaders can lower vaccine hesitancy in a country. Using a televised speech by France's President Emmanuel Macron as an exogenous shock, the results of this paper imply that such a form of health communication can significantly increase the willingness of individuals to get vaccinated against COVID-19 in a country. Thus, science-based health communication can be an effective weapon to change the beliefs of unvaccinated citizens. Both papers provide important policy implications for future pandemics and vaccination campaigns.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, *225*(2), 200–230.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

De Chaisemartin, C., & d'Haultfoeuille, X. (2022). *Difference-in-differences estimators of intertemporal treatment effects* (tech. rep.). National Bureau of Economic Research.

De los Santos, B., & Wildenbeest, M. R. (2017). E-book pricing and vertical restraints. *Quantitative Marketing and Economics*, *15*, 85–122.

Foros, Ø., Kind, H. J., & Shaffer, G. (2017). Apple's agency model and the role of most-favored-nation clauses. *The RAND Journal of Economics*, *48*(3), 673–703.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.

Huntington-Klein, N. (2021). *The effect: An introduction to research design and causality*. CRC Press.

Johnson, J. P. (2017). The agency model and mfn clauses. *The Review of Economic Studies*, *84*(3), 1151–1185.

Li, J., & Moul, C. C. (2015). Who should handle retail? vertical contracts, customer service, and social welfare in a chinese mobile phone market. *International Journal of Industrial Organization*, *39*, 29–43.

Lu, L. (2017). A comparison of the wholesale model and the agency model in differentiated markets. *Review of Industrial Organization*, *51*, 151–172.

Mallapaty, S. (2021). Can covid vaccines stop transmission? scientists race to find answers. *Nature*, *10*(10.1038).

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55.

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of educational Statistics*, *2*(1), 1–26.

Affidavit

I hereby declare that I completed the papers submitted and listed hereafter independently and with only those forms of support mentioned in the relevant paper or in the following supplementary list. When working with the authors listed, I contributed no less than a proportionate share of the work. In the analyses that I have conducted and to which I refer in the papers, I have followed the principles of good academic practice, as stated in the Statute of Justus Liebig University Giessen for Ensuring Good Scientific Practice.

Gießen, February 14, 2024

Phil-Adrian Klotz