Doctoral Thesis

Essays on the Industrial Organization of Transport, Retail and Health Markets

Jan Thomas Schäfer, M.A.

Submitted to Justus Liebig University Giessen,
Department of Economics

First supervisor: Prof. Dr. Georg Götz
Second supervisor: Prof. Dr. Peter Winker

December 12, 2024

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Georg Götz, for his invaluable guidance, encouragement, and support throughout my doctoral studies. His expertise, insightful feedback, and patient mentorship have been crucial to the development of this thesis. I am also grateful to my co-supervisor, Prof. Dr. Peter Winker, for their constructive criticism and continuous encouragement, which have significantly enriched my research.

This research was made possible by the generous funding provided by Deutsche Bahn AG, Börsenverein des Deutschen Buchhandels e. V., and Apothekerverband Westfalen-Lippe e.V., whose financial support allowed me to carry out my work without distractions.

I would like to acknowledge my colleagues and fellow doctoral candidates, especially Maximilian Maurice Gail, Samuel de Haas, Daniel Herold and Phil-Adrian Klotz, for their intellectual camaraderie, constructive discussions, and assistance with various aspects of my research. The collaborative environment at Justus Liebig University has been both inspiring and motivating.

On a personal note, I would like to thank my family for their unwavering support and patience throughout this challenging journey. To my parents, Dörte Hausrath-Schäfer and Hans Schäfer, your belief in me has been a constant source of strength. To my partner, Kseniya Fadeeva, thank you for your love, understanding, and encouragement. You have been my pillar throughout this entire process.

Lastly, I would like to extend my gratitude to my friends, whose support, humor, and understanding have kept me grounded throughout my doctoral journey.

Contents

1. Introduction

This doctoral thesis, *Essays on the Industrial Organization of Transport, Retail, and Health Markets*, investigates how regulatory changes, market structures, and government intervention shape outcomes in these sectors. The manuscripts examine topics including the regulation of Rx rebates in health markets, the strategic use of extended shopping hours in grocery retail, entry deterrence in transport markets, and the financial frameworks of European railways. Despite their diverse contexts, the manuscripts share a focus on analyzing the effects of regulatory changes and market adjustments within the framework of industrial organization. They employ similar methodological approaches, using variation across entities - such as states or regions - or exploiting (natural) experiments to identify causal effects. Most manuscripts rely on panel data and apply methods such as Data Envelopment Analysis (DEA), Fixed-Effects Models, Difference-in-Differences (DID), and Instrumental Variable techniques.

In the first manuscript, *Upstream vs. Downstream Grants - The Role of Public Contributions in Improving Railway Efficiency in Europe*, I investigate the effects of government funding structures on railway efficiency. The study uses a database established in Schäfer and Götz (2017) and further described in Götz and Schäfer (2020). These papers, coauthored by me, provide a detailed foundation for the dataset. Due to their more descriptive character, they are not included as part of this dissertation but have been placed in the appendix (see Appendix A.1 and A.2). Using Data Envelopment Analysis (DEA) and a second-stage truncated regression model (Simar & Wilson, 2007), the study examines whether public contributions are more effective when directed upstream to infrastructure managers or downstream to service providers. The results reveal that railways are more efficient when a larger share of government support is allocated to infrastructure managers, as this potentially lowers access charges, stimulates competition, and enhances network utilization. This work addresses gaps in the literature by explicitly analyzing the trade-off between upstream and downstream grants in a vertically separated railway system, a topic overlooked in prior studies that focused predominantly on total government support or the intensity of support compared to total cost (e.g., Cantos et al., 1999; Friederiszick et al., 2003; Oum & Yu, 1994). Furthermore, the analysis contributes to understanding the complex interdependence between infrastructure funding and public service compensation, highlighting how financial decisions at one level influence efficiency across the sector.

In the second manuscript, *Entry Deterrence Due to Brand Proliferation: Empirical Evidence from the German Interurban Bus Industry*, we analyze strategic behavior in the context of Flixbus's 2016 acquisition of Postbus. Using route-level data and employing Difference-in-Differences (DID) and Fixed-Effects regression models, the study investigates whether Flixbus engaged in pre-emptive brand proliferation to deter market entry. Our findings indicate that, prior to the acquisition, Flixbus offered a high frequency of trips on routes not served by Postbus, a strategy consistent with pre-emptive entry deterrence by product proliferation (Dixit, 1982; Schmalensee, 1978). Post-acquisition, the frequency of trips decreased significantly, and prices slightly declined, suggesting a strategic shift once competitive threats diminished. The study contributes to the literature on entry deterrence, which has primarily focused on industries such as airlines (Goolsbee & Syverson, 2008) and pharmaceuticals (Ellison & Ellison, 2011), by providing empirical evidence from the interurban bus market. Furthermore, it highlights the interaction between intramodal competition among bus operators and intermodal competition with railway services, consistent with findings on intermodal substitution (Beestermöller, 2017; Böckers et al., 2015). By examining the interconnected dynamics of frequency, pricing, and competition, this manuscript offers insights into how incumbents adapt strategies in highly differentiated markets with significant entry costs.

In the third manuscript, *Shopping Hours and Entry - An Empirical Analysis of Aldi's Opening Hours*, we examine Aldi's strategic decision to extend store hours past 8 p.m., interpreting this as entry into the market for late-night grocery shopping. Using a Logit regression model based on entry theory (Toivanen & Waterson, 2005) and a novel dataset covering the opening hours of nearly all German grocery retailers, we analyze factors that influence these decisions. While earlier theoretical work has focused on welfare implications and strategic differentiation of shopping hours (e.g., Inderst & Irmen, 2005; Shy & Stenbacka, 2006), empirical studies on this dimension of competition remain scarce. Our findings reveal that Aldi is more likely to extend hours if nearby Aldi outlets already stay open past 8 p.m., suggesting consumer learning and potential market expansion effects (Flores & Wenzel, 2016; Kosfeld, 2002). Conversely, Aldi is less likely to extend hours in areas where competitors such as Lidl and Netto close earlier, indicating that firm learning about low local demand also plays a role. These results highlight the strategic use of shopping hours to adapt to local market conditions and competitive pressures.

In the fourth manuscript, *The Substitutability Between Brick-and-Mortar Stores* *and e-Commerce - The Case of Books*, we analyze the interplay between physical bookstores and e-commerce in the German book market, a sector characterized by fixed book prices. This study addresses key gaps in the literature. While existing research extensively explores the rise of e-commerce and its competitive effects across industries (e.g., Goldmanis et al., 2010; Goolsbee, 2001), few studies examine markets where price competition is absent. By investigating the German book market, this paper provides insights into how non-price competition shapes the substitutability of online and offline retail channels. The manuscript also contributes to understanding the effects of digitization on traditional retailers. Prior research highlights the role of convenience, product variety, and service differences in driving consumer preferences (Brynjolfsson et al., 2003; Guo & Lai, 2017). Our findings reveal that physical bookstores and e-commerce are imperfect substitutes, as the closure of a bookstore leads to a significant decline in overall book sales, particularly for genres like fiction and children's books, where the offline channel complements online sales. While Brynjolfsson et al. (2006) emphasize the advantages of e-commerce in providing access to a wider variety of products and Reimers and Waldfogel (2021) highlight the benefits of crowd-sourced reviews for consumer decision-making, our results show that e-commerce does not fully compensate for the loss of services and discovery opportunities provided by physical bookstores. Lastly, the study contributes to policy discussions on vertical restraints such as resale price maintenance (RPM). Fixed book prices, common in Germany and other markets, are often justified as supporting cultural goods. Our results suggest that RPM can mitigate competitive pressures on brick-and-mortar bookstores but may not fully counteract the effects of digitization. By employing a novel dataset and an Instrumental Variable (IV) approach to address endogeneity, the study offers robust evidence on the economic implications of preserving physical retail infrastructure in the digital age.

In the fifth manuscript, *Efficiency in COVID-19 Vaccination Campaigns - A Comparison Across Germany's Federal States*, we analyze the efficiency of the COVID-19 vaccination rollout using Data Envelopment Analysis (DEA) and Fixed Effects panel regression. This study builds on the literature examining the logistical and organizational challenges of vaccination campaigns, which are critical to maximizing vaccine administration under constraints such as supply scarcity and the need for second-shot prioritization (ECDC, 2021; Mills & Salisbury, 2021). While previous studies have applied DEA to assess the performance of hospitals and vaccination centers (Staat, 2006; Valdmanis et al., 2003), this manuscript provides a novel application to the regional rollout of COVID-19 vaccines in Germany, focusing on the relationship between inputs (vaccine deliveries and reserves) and outputs (administered doses). The findings reveal significant efficiency differences across Germany's 16 federal states and highlight that integrating doctors' offices into the campaign on April 5, 2021, led to a structural break, increasing weekly vaccinations by 11.6 per 100 available doses. This complements existing evidence that decentralization and expanding the network of vaccine providers improve rollout efficiency (Ratzan et al., 2021). Additionally, the results underscore the importance of avoiding excessive vaccine reserves, as unused doses delay the campaign's progress and can prolong public health risks (Torres et al., 2021). By quantifying the potential gains from improved efficiency, this manuscript contributes to the broader discussion on optimizing vaccination strategies during pandemics. The study provides actionable insights for policymakers, demonstrating how systematic efficiency analyses can address regional disparities and enhance the effectiveness of public health interventions.

Finally, the last manuscript, *Evaluation of a Partial Ban of Rx-Rebates in Germany Using Difference-in-Differences*, advances the literature on price competition, digitization, and resale price maintenance (RPM) in the pharmaceutical sector. While previous research has primarily examined pharmacies' roles in providing primary care and over-the-counter medications (Agomo, 2012; Smith, 2009), this study shifts focus to how price competition impacts pharmacy profitability and financial sustainability, providing new insights into their role in public health. Additionally, the manuscript contributes to the debate on digitization, particularly the substitutability of online and offline services, by showing how rebates offered by online pharmacies can cannibalize the market share of brick-and-mortar pharmacies (Brynjolfsson & Smith, 2000; Cavallo, 2017). Using a difference-in-differences approach and a novel high-frequency dataset covering nearly half of all German pharmacies, the study finds that Germany's 2020 Local Pharmacy Support Act (VOASG) led to a modest increase in offline pharmacy sales (1.36-1.65%) but disproportionately benefited high-revenue pharmacies. Lastly, the findings extend research on RPM (Marvel & McCafferty, 1985; Telser, 1960), revealing that while the policy shifted profits toward traditional pharmacies, it failed to reverse the overall decline in pharmacy numbers, highlighting the limitations of RPM in achieving public health objectives in highly regulated markets.

2. Transport Markets

2.1. Upstream vs. downstream grants - The role of public contributions in improving railway efficiency in Europe

Author: Jan Thomas Schäfer

The manuscript has been presented at the following conference:

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

Upstream vs. downstream grants - The role of public contributions in improving railway efficiency in Europe

Jan Thomas Schäfer*

December 12, 2024

Abstract

The level of government support significantly influences the performance of European railways. However, prior analyses have largely focused on the sector as a whole, neglecting the distribution of public budget contributions between the upstream infrastructure manager and downstream service providers. This study employs a two-stage procedure involving Data Envelopment Analysis (DEA) in the first stage and a second-stage regression analysis to evaluate railway efficiency and analyze the relationship between funding structures and performance. Using a dataset covering eight European countries from 2001 to 2022, the results indicate that railways achieve higher efficiency when the upstream infrastructure manager receives a larger share of government funds, while downstream subsidies are relatively limited. Moreover, total operating contributions consistently enhance efficiency, whereas the impact of investment grants varies depending on the specification. These findings underscore the importance of balanced funding strategies that prioritize upstream contributions to foster competition and promote efficient use of public resources.

*Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen, Germany, E-mail: jan.t.schaefer@wirtschaft.uni-giessen.de.

1 Introduction

Research has shown that government support impacts the performance of European railways, explaining differences in productivity and efficiency between systems (e.g., Cantos et al., 1999; Friederiszick et al., 2003; Oum & Yu, 1994). However, existing studies primarily analyze the sector as a whole, neglecting the distribution of public budget contributions between upstream infrastructure managers and downstream transport operators.

European regulations permit various forms of government support, including investment grants, revenue contributions to infrastructure managers, and public service compensation for transport operators. These funding schemes are interlinked through the vertical structure of the industry, where access charges connect upstream and downstream levels. Adjustments in funding at one stage can influence the financial needs and incentives of the other, creating a trade-off in financing structures. For example, lower access charges resulting from upstream subsidies reduce public service compensation requirements and enhance competition. However, excessive reliance on tax-based financing can encourage inefficiencies, while user-based financing can undermine competitiveness. These considerations highlight the complexity of designing funding schemes that align with social, economic, and environmental goals.

This paper examines the impact of different funding structures on sector performance, with a focus on the trade-off between inefficiencies at the infrastructure level and those arising from public service subsidies. To explore this relationship, I use Data Envelopment Analysis (DEA) to measure efficiency and a second-stage truncated regression model (Simar & Wilson, 2007) to assess how funding structures influence performance. While DEA efficiency scores are calculated for a dataset covering nearly all European countries, the analysis of government grants' influence on performance is focused on eight countries: France, Germany, Italy, Norway, Spain, Sweden, Switzerland, and the United Kingdom. The study spans data from 2001 to 2022.

The results indicate that railway systems achieve higher efficiency when the upstream infrastructure manager receives a larger share of government funds, while downstream subsidies to service providers remain limited. These findings suggest that restructuring funding schemes could foster competition and more efficient use of public resources, provided that robust regulatory frameworks are in place.

This paper contributes to the extensive literature on the performance of European railways, which has been a focus of research for decades. Recent studies have examined the effects of competition and vertical separation on efficiency (e.g., Álvarez-SanJaime et al., 2024; Fitzová, 2022) and explored regional disparities (Fitzová & Nash, 2024). The role of direct government support in shaping sector performance has also been analyzed in several studies. However, only a few, such as Cantos et al. (1999), Friederiszick et al. (2003), and Oum and Yu (1994), have employed advanced performance measurement methods like Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA). Other authors have relied on Partial Productivity Measures (PPM) to link performance indicators with sector financing.

Some recent analyses lack scientific rigor. For instance, the Boston Consulting Group (2012) developed a weighted performance index (RPI) that combines indicators such as network utilization intensity, service quality, the share of high-speed rail, and safety aspects. Their findings indicate a positive correlation between public costs - defined as the sum of government contributions for operations, maintenance, and investments - and performance. By contrast, Friederiszick et al. (2003) used SFA to explore the impact of state aid on efficiency more comprehensively. They found that while state aid has a positive effect on efficiency, its intensity (state aid as a proportion of total operating costs) negatively influences efficiency scores. Similarly, Oum and Yu (1994) showed that higher levels of government support relative to operating costs adversely affect efficiency. Their DEA second-stage Tobit regression suggests that railways receiving significant subsidies tend to be less efficient than their counterparts. Likewise, Cantos et al. (1999) find that railways are more efficient when they are less dependent on public subsidies.

Panel data studies, such as Laabsch and Sanner (2012), have examined public contributions' effects on modal share in passenger rail transport. They concluded that higher public contributions do not significantly influence modal share but may correlate with higher sector costs, potentially negating the benefits of the subsidies. Other researchers, like Kyriacou et al. (2019), have incorporated investments (as a share of GDP) as input variables in DEA models to highlight the importance of government quality in ensuring efficient investments.

The trade-off between upstream and downstream subsidization has been addressed in a limited number of studies, most of which focus on climate or trade policy (Fischer et al., 2014; Hokari et al., 2003). While upstream subsidies are prevalent in network industries, downstream subsidies have been primarily examined in trade policy contexts (e.g., Bernhofen, 1997; Hamilton & Requate, 2004). Thus, this paper contributes to the literature on upstream and downstream subsidization by presenting evidence from the transport sector.

The remainder of the paper is organized as follows. Section 2 provides an overview of the organizational structure of railways in European countries and the principles of their financing. Section 3 details the methodology employed for the empirical analysis. Section 4 describes the data sources and presents descriptive statistics on government grants to European railways. The results of the first-stage DEA analysis are presented in Section 5.1, while Section 5.2 discusses the empirical assessment of the influence of government grants on performance. Finally, Section 6 concludes the paper.

2 Financing of Railways in Europe

Government support plays a major role in the financing of the European railway sector. Typically infrastructure managers, regional transport operators as well as some freight and long-distance passenger rail operators receive monetary contributions from public budgets to supply the market with services and to undertake investments in rolling stock and infrastructure. Figure 1 shows a simplified illustration of the organizational structure of railways in European countries, it also includes relevant directives and regulations that set out the framework for government support and track access charges.

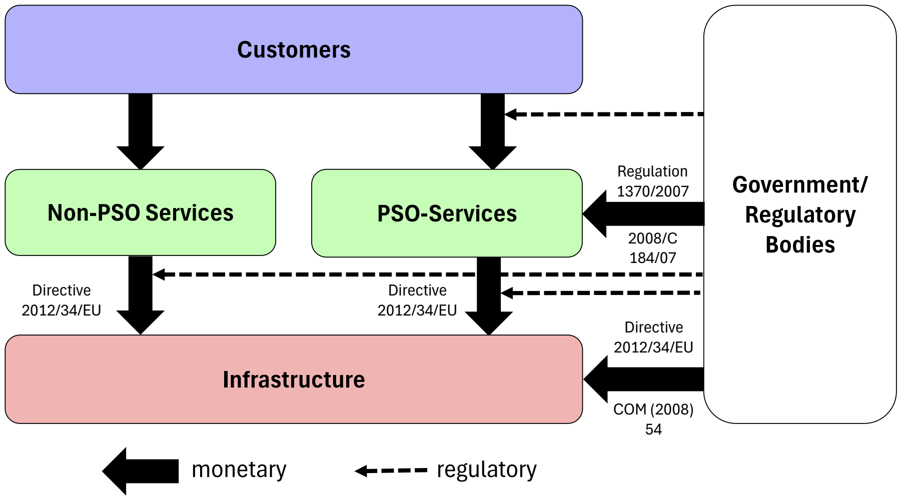


Figure 1: Overview of revenue sources and regulatory frameworks for European railway services, highlighting contributions from customers, non-PSO services, government PSO services, and infrastructure. Key regulations, including Directive 2012/34/EU and Regulation 1370/2007, are indicated as guiding frameworks.

Transport services are typically divided into two categories: services that are subject to a Public Service Obligation (PSO) and services that are operated on a commercial basis. The former usually refers to regional public passenger transport services. However, in some countries, long distance passenger transport services are subject to a Public Service Obligation. Non-PSO services include all other passenger transport services, i.e. services that are run on a commercial basis, as well as freight transport services. These services are run without government funding. However, as pointed out in the community guidelines on State aid for railway undertaking (2008/C 184/07), commercial operators might also receive funding, in particular for the purchase of rolling stock.

Public service operators, generally speaking, receive a compensation depending on the costs incurred, minus the receipts from tariff plus a reasonable profit (Regulation 1370/2007). I will refer to these payments using the term public service compensation (PSC). The level of necessary public service compensation is determined by the level of tariff receipts as well as the cost of services. In most cases, tariffs for PSO services are set by the competent authority. Costs of public services (and other operators) depend to a large extend on the level of payments for the use of the network, i.e. access charges. Access charges are regulated according to Directive 2012/34/EU and should be set on the level of direct costs plus a mark-up to obtain full cost recovery. The level of mark-ups crucially depends on the level of fixed costs as well as the amount of government financing that the infrastructure manager receives for the operation of the network. According to Communication COM (2008) 54 of the European Commission payments for operating the network are likely to supplement user charges. Therefore, I will refer to them using the term revenue contributions.

While public service compensation and infrastructure revenue contributions mainly substitute user-based financing, public funding of investments mainly impacts the necessary amount of loans to be raised by the infrastructure manager. While in some countries infrastructure investments are mainly financed from public funds, other countries let the infrastructure manager get into debt. This usually means, that financing costs, i.e. interest expenses, are higher. Some governments compensate for the additional burden, while in other countries the market needs to bear the additional financing costs.

If revenue contributions cover a large share of operating expenses of the infrastructure manager, mark-ups will be low. Thus, the overall level of access charges will be lower. Consequently, transport service providers need to pay less for the use of the infrastructure. In return, less public service compensation need to be payed to compensate companies for the provision of services. Both services subject to a Public Service Obligation and those not subject to it might pass cost reductions through to consumers (Arrigo & Di Foggia, 2013). I will investigate the impact of this mechanism on the performance of the sector in the following chapters.

3 Methodology

3.1 Methods of Performance Measurement in the Railway Sector

The performance of railway systems has been extensively studied (Catalano et al., 2019; Oum et al., 1999). By comparing indicators across countries, sectors, or entities, these studies aim to identify best practices and opportunities for improvement. Economists use performance measures to analyze railway reforms, their impact on efficiency, and the influence of different operating environments.

Performance encompasses several interrelated dimensions: efficiency, productivity, and effectiveness. Efficiency assesses the optimal use of resources, focusing on input minimization or output maximization. Productivity measures the output produced relative to inputs used. Effectiveness evaluates goal achievement, often relative to resource utilization. Five primary methods of performance measurement have been applied in the railway sector: partial productivity measures, total factor productivity, data envelopment analysis (DEA), stochastic frontier analysis (SFA), and (corrected) ordinary least squares estimation (Catalano et al., 2019; Merkert et al., 2010; Oum et al., 1999).

DEA is the most widely used efficiency method in recent railway and transport studies, followed by SFA (Catalano et al., 2019; Cavaignac & Petiot, 2017). DEA, a non-parametric approach based on linear programming, compares the efficiency of decision-making units (DMUs) such as railway sectors, train operators, or infrastructure managers. Using input and output data, DEA constructs a piecewise linear production frontier. Efficiency is determined by measuring the distance of each DMU from this frontier. DEA can be input- or output-oriented, depending on whether the analysis minimizes input for a given output or maximizes output for a given input. When price data is available, cost or revenue frontiers can also be derived (e.g., Cantos et al., 2002).

SFA, in contrast, is a parametric approach that uses maximum likelihood estimation to model a stochastic production frontier. Like DEA, SFA can estimate cost frontiers reflecting the minimum cost of producing a given output at specific input prices (Holmgren, 2013). However, SFA requires assumptions about the functional form of the production function and the distribution of inefficiency, which can lead to biased estimates if incorrectly specified. DEA avoids these functional form assumptions but is sensitive to the choice of scale (constant vs. variable returns to scale). Additionally, DEA lacks inherent statistical tests for sensitivity, making SFA advantageous for some applications (e.g., Cantos et al., 2012; Fiorentino et al., 2006). Corrected ordinary least squares (COLS) is another approach occasionally used to estimate deterministic production frontiers, although DEA and SFA dominate the literature (Coelli & Perelman, 1996; Coelli et al., 2005).

Partial productivity measures (PPM), also known as partial factor productivity, are simpler tools that relate a single input to a single output, providing insights into sectoral productivity. Examples include the operating ratio (operating costs/operating revenue), revenue per traffic unit, traffic units per employee, and traffic density (traffic units per track-km) (Beck et al., 2013; NERA, 2004; Oum et al., 1999; World Bank, 2011). Efficient railway systems achieve higher asset utilization and output relative to costs under comparable conditions. In contrast to PPM, total factor productivity (TFP) aggregates inputs and outputs, offering a broader performance perspective (Tretheway et al., 1997).

This study uses DEA as the primary method for performance evaluation due to its flexibility and data-driven nature. Unlike parametric approaches such as SFA or COLS, DEA does not require assumptions about the functional form of the production function, making it particularly suitable for analyzing heterogeneous railway systems with varying operating environments. DEA's ability to accommodate multiple inputs and outputs simultaneously is advantageous when evaluating complex systems like railways, where performance depends on diverse factors. While SFA incorporates stochastic noise, the deterministic framework of DEA is sufficient for this analysis given the quality and consistency of the available data.

3.2 Two-Stage Procedure for Evaluating Railway Efficiency

I apply the two-stage procedure proposed by Simar and Wilson (2007) to analyze the impact of government funding on the performance of European railways. In the first stage, DEA efficiency scores are calculated. The second stage comprises a bootstrap based two-stage estimation that yields estimated standard errors and confidence intervals that do not suffer from bias due to estimated efficiency scores being correlated. It also overcomes issues of sample selection (efficiency scores are calculated from a common sample of data).

First-Stage: DEA

First, efficiency scores are estimated. The output-oriented DEA with constant returns to scale, can be formalized as follows (see Charnes et al. 1978). Given n decision-making units (DMUs), each using m inputs $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m_+$ to produce s outputs $\mathbf{y} = (y_1, \ldots, y_s) \in \mathbb{R}^s_+$, the output-oriented DEA with constant returns to scale (CRS) is obtained by solving the following linear programming problem, for each DMU i:

$$\max \quad \phi_i$$

subject to

$$\sum_{j=1}^{n} \lambda_j x_{k,j} \leq x_{k,i} \quad k = 1, \ldots, m,$$

$$\sum_{j=1}^{n} \lambda_j y_{r,j} \geq \phi_i \cdot y_{r,i} \quad r = 1, \ldots, s,$$

$$\lambda_j \geq 0 \quad j = 1, \ldots, n,$$

where ϕ is the scalar that reflects the proportional increase in outputs achievable by DMU i without changing the input levels, and $\lambda = (\lambda_1, \ldots, \lambda_n)$ are the intensity variables (weights for other DMUs in the reference set). A score of $\phi = 1$ indicates that the DMU is on the efficient frontier, while $\phi > 1$ suggests inefficiency. In the case of variable returns to scale (VRS), an additional constraint $\sum_{j=1}^{n} \lambda_j = 1$ is added, which ensures that each DMU is compared only to convex combinations of other DMUs, thus allowing for VRS.

For the orientation and assumption of returns to scale, I adopt an output-oriented model with constant returns to scale (CRS). This choice is appropriate because the railway network length, one of the inputs, is relatively fixed and cannot be significantly adjusted between periods. Therefore, it is more plausible to assume that inputs remain constant while outputs adjust. Previous research indicates that the choice of orientation is generally not critical for railways (e.g., Coelli & Perelman, 2000). The assumption regarding returns to scale appears to be more critical. However, repeating the analysis under the assumption of variable returns to scale (VRS) produces similar, yet less robust, results (see Table B.4 for the second-stage results).[1]

[1] Note that the VRS model can only be estimated for DEA efficiency scores calculated using a pooled frontier, as countries in the smaller subset are on the yearly frontiers in many

Since the data used in this study has a panel structure, different approaches for the computation of DEA efficiency scores are employed in the literature. Some authors, such as Oum and Yu (1994), Coelli and Perelman (1999), and Kleinová (2016), treat each observation, i.e., each combination of time and country, as an independent decision-making unit (DMU). Thus, they measure the efficiency of each DMU against a pooled frontier. The resulting efficiency scores capture both technical progress and catching-up effects (see Growitsch and Wetzel 2007). Other studies run separate DEA models for each period, while some apply the Malmquist Index to decompose productivity changes over time into efficiency change (catching-up) and technical change (frontier shift) components (see Färe et al. 2011 for technical details). Additionally, some authors model dynamic DEA frameworks, for example, by accounting for the fact that today's (quasi-fixed) inputs are yesterday's outputs (see Lim et al. 2022).

As the main specification, I calculate an independent efficiency frontier for each year, rather than relying on a pooled frontier across the entire period. This approach ensures that efficiency scores are measured relative to contemporaneous peers, accounting for temporal variations in economic conditions, regulatory changes, and external shocks. For example, the COVID-19 pandemic likely disrupted productivity and operating conditions, making it inappropriate to compare pandemic-era performance with other years. Year-specific frontiers provide a fairer assessment of efficiency by isolating performance within comparable periods.

years. Since these observations are dropped from the regression due to truncation, there are too few observations for the maximum likelihood estimation to converge. Therefore, these results cannot be directly compared to those of the main specification.

Although year-specific frontiers reduce the sample size for each year's estimation, this limitation is outweighed by their advantages. A pooled frontier assumes stable production technology and environmental conditions over time, which may not hold in dynamic contexts. By adopting yearly frontiers, I avoid the potential biases of pooling data across periods with differing production environments, ensuring a more accurate comparison of efficiency scores across countries. However, repeating the analysis for a pooled frontier, yields similar results (see Table B.2 for the DEA efficiency scores and Table B.3 for the second-stage results).

Second-Stage: Two-stage estimation

In the second-stage of the analysis, DEA efficiency scores are regressed on several explanatory variables. Standard OLS is unsuitable here due to boundary constraints and potential serial correlation among the efficiency estimates. These boundary issues arise because DEA scores are bounded - typically, efficiency scores lie between zero and one, depending on whether inefficiency is defined from above or below. Simar and Wilson (2007) address these challenges by proposing two algorithms that account for both, the bounded nature of DEA scores and the complex correlation structure among them. Both algorithms involve bootstrapping to enhance the accuracy and reliability of efficiency estimates. In this study, I will use what Simar and Wilson (2007) call *Algorithm 1.*

The steps in this approach involve: (a) conducting a truncated regression of the DEA scores on the explanatory variables, which respects the bounded nature of the efficiency estimates; (b) generating artificial efficiency scores through bootstrapping to account for sampling variability; and (c) calculating bias-corrected efficiency scores, along with bootstrap standard errors and confidence intervals. This multi-step process improves the reliability of inference by adjusting for both truncation and potential bias, making it a robust approach for second-stage DEA analysis. The feasibility and effectiveness of this method are further supported in Simar and Wilson (2011).

The structural equation of the second-stage model has the following form:

$$\phi_i = \mathbf{X}'_{i,t}\beta + \epsilon_{i,t} \tag{1}$$

where $\phi_{i,t}$ represents the technical efficiency score of DMU i in year t, $\mathbf{X}_{i,t}$ is a vector of exogenous (environmental) variables, β is a vector of coefficients, and $\epsilon_{i,t}$ is the error term.

To estimate the structural Equation (1), *Algorithm 1* from Simar and Wilson (2007) follows a structured bootstrapping process. First, the initial DEA efficiency scores $\phi_{i,t}$ for which $\phi_{i,t} > 1$ holds are included in a truncated regression (left-truncated at 1) of $\phi_{i,i}$ on the exogenous variables $\mathbf{X}_{i,t}$. This step provides initial estimates for the coefficients, β, and the variance parameter, σ, through maximum likelihood estimation.

Next, a bootstrap procedure is applied, consisting of the following steps, repeated B times to obtain a set of B bootstrap estimates $(\beta^{(b)}, \sigma^{(b)})$ for $b = 1, \ldots, B$:

1. For each observation (i, t), an artificial error term $\epsilon_{i,t}^{(e)}$ is drawn from a truncated normal distribution $N(0, \sigma)$, truncated at $1 - \mathbf{X}'_{i,t}\beta$.

2. Artificial efficiency scores $\phi_{i,t}^{(e)}$ are then generated as $\phi_{i,t}^{(e)} = \mathbf{X}'_{i,t}\beta + \epsilon_{i,t}^{(e)}$ for each observation (i, t).

3. A truncated regression (left-truncated at 1) of $\phi_{i,t}^{(e)}$ on $\mathbf{X}_{i,t}$ is conducted to obtain maximum likelihood estimates $\beta^{(b)}$ and $\sigma^{(b)}$ for each bootstrap iteration.

After completing B iterations, the bootstrap distributions of β and σ are used to compute confidence intervals and standard errors, providing robust inference for the second-stage DEA model.

4 Data Sources and Descriptive Statistics

The variables used as input or output measures in the first-stage vary across existing studies and are influenced by the data sources employed. Some researchers utilize data from the UIC, available at the company level (see Fitzová and Nash 2024 and Álvarez-SanJaime et al. 2024). However, UIC data primarily includes major (incumbent) railway undertakings, potentially limiting the scope of analyses. Other authors rely on country-level data, which can be obtained from sources such as Eurostat, the International Transport Forum, or the World Bank (see Kleinová 2016 for an application using country-level data). It is not always clear, however, to what extent the country-level data includes smaller undertakings.[2] Some authors combine information from the UIC with country-level data (see, for example, Niu et al.

[2] In Germany, for example, only freight companies with a transport volume exceeding 10 million tonne-kilometers are legally required to report data to the national statistical service.

2023). In the following analyses, I will mainly use country-level data from the Statistical Pocketbook (see European Commission and Directorate-General for Mobility and Transport 2024).

Regarding the data on revenue contributions and public service compensation for the second-stage, I will use country-level data from an updated version of the dataset compiled by Schäfer and Götz (2017). Data sources include annual reports from infrastructure managers, railway undertakings, and public budget reports. However, the scope of the data limits the analyses in the second-stage because detailed data on revenue contributions and public service compensation for European railways are not centrally collected by European or transnational institutions. The updated version of the dataset from Schäfer and Götz (2017) covers eight European countries, namely France, Germany, Italy, Norway, Spain, Sweden, Switzerland, and the United Kingdom, spanning the period from 2001 to 2022.[3]

Figure 2 shows a map with the countries for which data is available. Countries that enter the first-stage are colored grey. Countries for which financing data is available, and that thus enter the second-stage truncated regression, are marked with a white hatch.

Table 1 contains summary statistics for the variables used in the DEA and the second-stage regression. The smallest network in the sample has a length of 204.2 kilometers, while the largest network has a total length of 41 531 kilometers. On average, 13.6 billion passenger-kilometers (*rail_pkm*) and 18.41 billion tonne-kilometers (*rail_tkm*) are produced. The average share of freight

[3]Data for Italy and Sweden is available only for certain years.

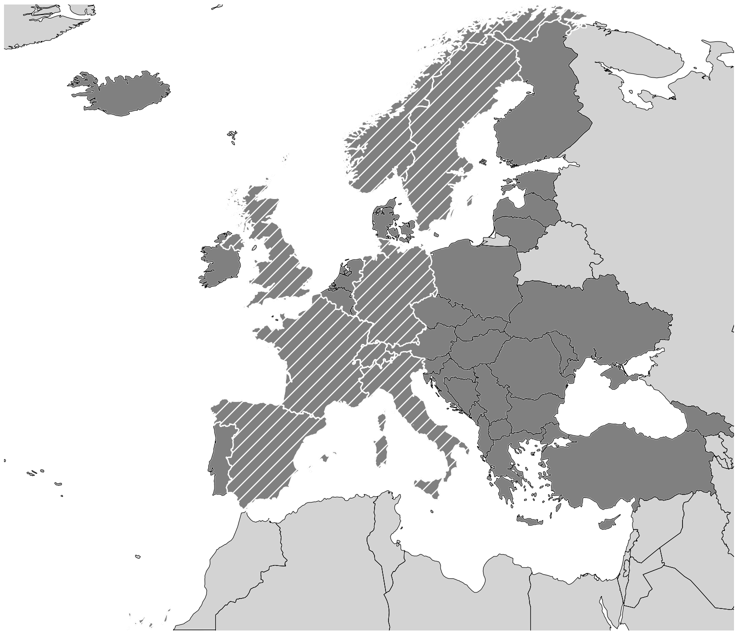


Figure 2: Countries in the sample

transport (*share_freight*), i.e., *rail_tkm* divided by the sum of *rail_pkm* and *rail_tkm* (*rail_ptkm*), is 59.83 %.

On average, an annual investment grant (*infra_inv_grant_pps_capita*) of € 71.01 per capita is granted to the infrastructure manager. The average revenue contribution to the infrastructure manager (*infra_rev_grant_pps_ptkm*) is € 0.023 per passenger-tonne-kilometer, while the average public service compensation amounts to € 0.025 per passenger-tonne-kilometer (*psc_pps_ptkm*). All monetary values are expressed in terms of Purchasing Power Standards (PPS), with the EU27 average price level, based on actual individual consumption, serving as the constant reference (EU27 = € 1).[4]

[4]Figures expressed in nominal values are presented in the appendix (Figure A.1 and Figure A.2).

Table 1: Summary statistics

	N	Mean	SD	Min	Median	Max
Inputs						
rail_length	763	7 303	8 584	204.2	3 631	41 531
rolling_stock_pocketbook	801	26 190	41 487	62	10 551	266 002
staff_uic	718	44 057	70 758	360	18 109	377 511
Outputs						
rail_pkm	769	13.6	22.61	0.000437	3.957	108.7
rail_tkm	778	18.41	38.02	0.009	8.526	262.5
Non-financial control variables						
pop_density	780	157.6	233.7	2.8	100.2	1 693
gdp_pps_capita	805	24 869	13 424	3 600	23 500	90 600
share_freight	762	0.5983	0.2456	0.02914	0.6431	0.9866
share_passenger	762	0.4017	0.2456	0.01341	0.3569	0.9709
Fincancial control variables						
infra_rev_grant_pps_ptkm	171	0.02297	0.02572	-2.3e-06	0.01355	0.1541
infra_rev_grant_pps_capita	171	34.28	29.75	-0.001926	22.2	115.7
infra_inv_grant_pps_ptkm	171	0.03956	0.03485	0	0.02666	0.1778
infra_inv_grant_pps_capita	171	71.01	63.13	0	51.79	246.3
psc_pps_ptkm	167	0.02531	0.02412	-0.007626	0.0229	0.236
psc_pps_capita	167	43.78	31.61	-10.09	35.1	139.8

The evolution of financial control variables over time, expressed on a per capita basis, is shown in Figure 3. These figures reveal notable differences across countries both in terms of levels and trends. Across all countries, infrastructure investment grants (*infra_inv_grant_pps_capita*) have steadily increased over the observed period. However, Spain exhibits a contrasting trend in investment grants, which show a significant decline over time, particularly after the onset of the financial crisis.

In 2022, there were notable differences in per capita grants for infrastructure investments across countries. Switzerland (CH) and Norway (NO) had the highest levels, with € 246.3 and € 240.1 per capita, respectively. These amounts were markedly higher than those of other countries in the sample.

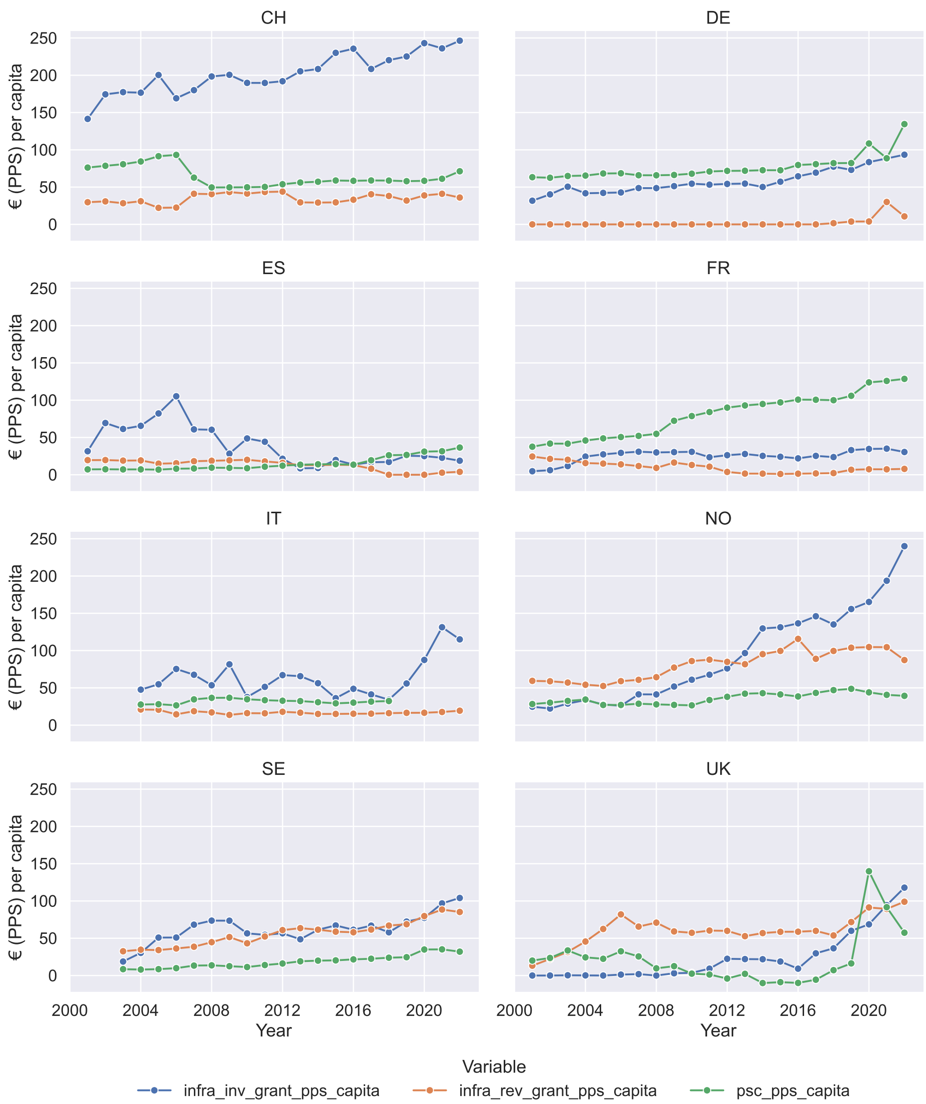


Figure 3: Grants for infrastructure investment, the operation of the infrastructure, and the provision of public services in € (PPS) per capita.

Italy (IT) followed with € 115.0 per capita, while the United Kingdom (UK) recorded € 117.9 per capita. Sweden (SE) had a slightly lower level of € 104.0

20

per capita, and Germany (DE) reached € 93.4 per capita. France (FR) had a lower level of € 30.5 per capita, and Spain (ES) recorded the lowest grants for investments in infrastructure at € 18.8 per capita.

The development of grants for the operation of infrastructure (*infra_rev_grant_pps_capita*) and for the provision of public services (*psc_pps_capita*) confirms the findings of Schäfer and Götz (2017) and Götz and Schäfer (2020). In general, two financing models are applied: either focusing on grants to upstream or downstream undertakings. Germany focuses government contributions on the operation of transport services, while the UK emphasizes revenue contributions to the infrastructure manager. Other countries apply a hybrid model, where both public service compensation and revenue contributions are granted - i.e., upstream and downstream undertakings receive almost equal shares of government grants.

However, with the onset of the pandemic, changes in the funding schemes of Germany and the UK can be observed. To support specific transport services, the German government allocated funds to reduce access charges. At the same time, the UK suspended rail franchising to maintain service as passenger demand fell due to the COVID-19 pandemic. Subsequently, it was decided to permanently abolish the rail franchising policy, effectively converting the franchises into concessions. This change had a notable effect on the level of public service obligations. Note that the UK publishes figures in terms of financial years, ranging from April to March. I attribute the figures to the year in which most months are situated; for example, the financial year 2020–2021 (ranging from April 2020 to March 2021) is shown as 2020 in the figures.

In the course of the war in Ukraine, Germany heavily subsidized public transport by introducing a nationwide flat-fee ticket (€ 9 per month) in 2022. To compensate for the forgone revenue, public service obligations were increased. In 2022, Germany's public service obligations amounted to € 134.5 per capita, representing one of the highest levels among the countries analyzed. France followed closely with € 128.6 per capita. Switzerland and the United Kingdom had notably lower levels of public service obligations, at € 71.2 and € 57.4 per capita, respectively. In contrast, Spain, Norway, and Sweden allocated even less, with € 36.5, € 39.3, and € 32.0 per capita, respectively. Data for Italy were not available for 2022, as Trenitalia stopped publishing annual reports.

At the same time, infrastructure revenue grants in Germany were among the lowest in 2022, amounting to € 10.7 per capita. This level was significantly lower compared to most other countries in the sample. For instance, Norway and Sweden had the highest levels of infrastructure revenue grants, at € 87.3 and € 85.2 per capita, respectively. The United Kingdom followed with € 99.0 per capita, reflecting its emphasis on revenue contributions to the infrastructure manager. Switzerland also reported a higher level, at € 36.0 per capita. France and Spain had relatively low levels, at € 7.9 and € 3.9 per capita, respectively, while Italy recorded € 19.4 per capita.

Figure 4 shows the evolution of financial control variables over time, expressed per passenger-tonne-kilometer (PTKM). Investment grants, infrastructure revenue grants, and public service compensations exhibit similar trends to those presented on a per capita basis. However, the reduction in transport volumes during the pandemic accentuates the effects of increased grants aimed at mitigating the crisis. In some cases, grants that declined between 2019 and 2020 on a per capita basis increased when measured per passenger-tonne-kilometer, due to the sharp decline in transport volumes. Since transport volumes had not returned to pre-crisis levels by the end of 2022 in all of the studied countries, the level of grants may still reflect the ongoing impact of the pandemic.

In 2022, infrastructure revenue grants were € 0.004 per ptkm in Germany, among the lowest in the sample. France followed with € 0.0037 per ptkm, while Spain recorded slightly higher levels at € 0.0049 per ptkm. Italy exhibited a more substantial level of infrastructure revenue grants at € 0.016 per ptkm, whereas Sweden and Norway had significantly higher levels at € 0.025 and € 0.065 per ptkm, respectively. The United Kingdom reported the highest value in the sample, with infrastructure revenue grants reaching € 0.099 per ptkm. Switzerland, despite its high overall infrastructure investment, recorded € 0.010 per ptkm.

Grants for the provision of public services were € 0.051 per PTKM in Germany in 2022. France recorded the highest level in the sample, with € 0.061 per ptkm, followed closely by the United Kingdom at € 0.058 per ptkm. Spain also demonstrated significant expenditure in this category, reaching € 0.045 per ptkm. In contrast, Norway and Switzerland allocated € 0.029 and € 0.020 per ptkm, respectively, while Sweden reported the lowest value at € 0.009 per ptkm.

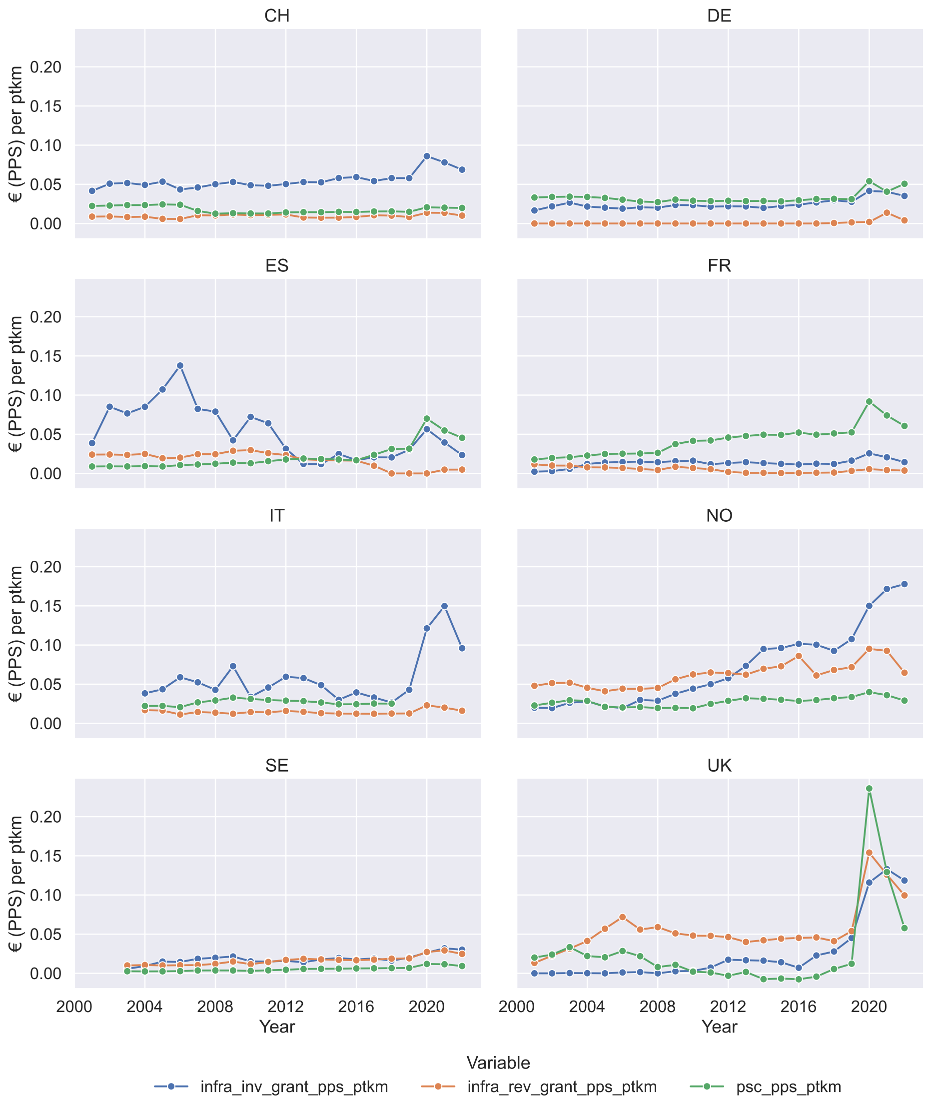


Figure 4: Grants for infrastructure investment, the operation of the infrastructure and for the provision of public services in € (PPS) per passenger-tonne-kilometer

As described before, countries with lower infrastructure revenue grants, such as Germany, Spain, and France, tended to allocate more funds per PTKM

24

toward public service compensations. This trend is particularly evident when comparing the United Kingdom, which had the highest infrastructure revenue grants at € 0.099 per PTKM, with Sweden, where low public service compensations coincided with relatively higher infrastructure revenue grants. The ratio of infrastructure revenue grants to public service compensations further highlights these differences: Germany, France, and Spain had ratios of 0.08, 0.06, and 0.11, respectively, reflecting a stronger reliance on public service compensations. In contrast, Sweden and Norway had significantly higher ratios, at 2.66 and 2.22, indicating a greater emphasis on infrastructure revenue grants. The United Kingdom exhibited a ratio of 1.72, whereas Switzerland had a ratio of 0.50, indicating that public service compensations were about twice as high as infrastructure revenue contributions.

In the following section, I will first compute DEA efficiency scores and then, in a subsequent second-stage regression, use the ratio of infrastructure revenue grants to grants for the provision of public services to explain differences in the efficiency scores.

5 Results

5.1 First-Stage: DEA Efficiency Scores

Table 2 presents the DEA efficiency scores for the output-oriented model with constant returns to scale, using the length of the rail infrastructure (*rail_length*), the size of the rolling stock (*rolling_stock_pocketbook*), and the average annual staff size (*staff_uic*) as input variables. Passenger-kilometers (*rail_* *pkm*) and tonne-kilometers (*rail_tkm*) are used as output variables. Values close to or equal to one indicate efficient decision-making units.

Several countries achieved efficient outcomes in one or more years, including Estonia, Latvia, the Netherlands, Norway, Sweden and Ukraine, which consistently exhibited efficiency scores of 1 over multiple years. These countries often have a high share of freight transport on their networks, a factor that may contribute to their efficiency. Switzerland and Denmark also reached efficient outcomes in specific years, particularly in 2020 and 2021, respectively. In contrast, countries such as Germany and France exhibited scores close to, but not at, full efficiency over most of the observed period, reflecting their relatively stable, albeit slightly suboptimal, performance. Meanwhile, missing data (NA) for several countries underscore the challenges of consistent data collection. Notably, some countries experienced significant inefficiencies, with scores well above 1, suggesting room for substantial improvement in their operations.

Table 2: Efficiency scores (yearly frontiers) - output oriented model, constant returns to scale

country	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
AL	1	1.55	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
AT	1.83	1.79	1.7	1.73	1.76	1.52	1.58	1.52	1.81	1.65	1.71	1.85	1.79	1.68	1.56	1.51	1.43	1.4	1	1.31	1.24	1
BA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6.94	7.24	5.03	6.83	8.17	4.7	4.51	3.85	3.27
BE	1.66	1.67	1.67	1.65	1.74	1.66	1.61	1.53	1.72	1.58	NA	NA	1	1.25	1.1	1.24	1.27	1.27	1.14	1	1.08	1.14
BG	4.11	4.68	4.24	4.58	4.68	4.27	4.59	5.09	6.85	6.65	7.22	8.44	4.35	4.02	3.6	5.95	3.48	3.77	3.68	2.97	2.84	2.21
CH	1.27	1.2	1.18	1.25	1.43	1.4	1.46	1.41	1.41	1.32	1.38	1.46	1.48	1.45	1.33	1.34	1.36	1.21	1.15	1	1	1.18
CZ	3.14	3.63	3.37	3.71	3.83	3.42	3.5	3.73	4.12	3.66	4.38	4.36	4	3.76	3.45	2.93	2.63	2.64	2.52	2.07	2.1	2.04
DE	1.34	1.57	1.74	1.84	1.83	1.67	1.66	1.62	1.81	1.68	1.8	1.86	1.79	1.79	1.64	1.63	1.54	1.56	1.49	1.38	1.38	1.06
DK	1.28	1.3	1	1.48	1.93	2.12	2.16	1.35	1.58	1.47	1.53	1.51	1.61	1.86	1.79	1.59	1.43	1.37	1.39	1	1	1
EE	1	1	1	1	1	1	1	1	1	1	1	1	NA	NA	1.54	1.75	1.74	NA	NA	NA	NA	NA
ES	1.67	1.61	1.69	1.92	2.04	1.96	1.92	1.66	1.69	1.87	2.53	2.86	3	2.82	2.08	1.85	1.79	1.73	1.66	1.67	1.31	1.25
FI	1.65	1.56	1.55	1.59	1.61	1.42	1.64	1.45	1.66	1.82	2.04	2.09	2.15	1.82	1.79	1.56	1.39	1.5	1.81	1.47	1.43	1.88
FR	1.47	1.4	1.43	1.52	1.93	1.96	2.02	1.89	1.91	1.86	1.86	1.9	2.01	2.03	1.77	1.79	1.72	1.37	1.31	1	1	1
HR	5.57	5.11	4.87	5.18	5.4	4.25	4.17	4.27	4.39	4.63	5.42	6.07	6.21	5.69	5.46	5.45	3.08	3.59	4.36	2.53	2.57	2.14
HU	2.97	2.64	2.73	2.87	3.24	3.25	3.38	2.95	3.22	2.82	3.26	3.11	2.68	2.46	2.4	3.22	1.96	2.37	2.31	2	2.04	1.87
IE	2.71	2.47	3.2	3.67	3.96	3.99	3.25	3.08	2.73	3.23	3.04	2.5	2.66	2.36	2.15	NA	1.82	1.5	1.47	1.9	1.96	NA
IT	1.38	1.28	1.3	1.39	1.69	1.75	1.83	1.83	1.94	1.94	2.02	2.05	2.1	1.99	1.83	1.81	1.54	1.32	1.31	1.39	1.08	1.02
LT	1.81	1.67	1.3	1.34	1.29	1.08	1	1.04	1.38	1.14	1.29	1.35	1.31	1.18	1.16	1.01	NA	1	1	1	1	NA
LU	2.28	2.78	2.89	3.07	3.57	3.11	2.61	3.65	4.11	3.32	3.75	3.96	NA	NA	NA	3.67	3.5	3.55	3.63	3.05	3.18	3.88
LV	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1.04	1.68	1.68	NA
MD	NA	NA	NA	NA	NA	NA	NA	NA	7.18	7.39	7.71	10.02	8.2	8.66	10.29	11.61	9.17	10.23	9.96	NA	NA	3.82
ME	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	4.61	6.59	7.21	6.26	5.85	NA
MK	1	1.47	1.85	8.02	6.68	5.19	4.38	4.82	5.89	4.65	5.4	7.02	6.33	NA	NA	13.34	10.25	10.5	8.92	8.86	7.82	7.12
NL	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
NO	1.41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	NA	NA	NA	NA	NA
PL	2.38	2.58	2.31	2.36	2.59	2.27	2.3	2.43	2.95	2.57	2.88	3.1	2.77	2.64	1.67	1.51	1.37	1.8	1.88	1.53	1.43	1.31
PT	2.05	1.92	2.27	2.44	2.6	2.62	2.46	2.14	1.47	2.35	2.66	2.78	2.92	2.65	1.58	2.04	1.76	1.77	1.76	1.44	1.55	1.61
RO	3.11	3.78	3.63	3.37	3.85	3.72	3.93	4.2	5.58	5.06	5.38	5.72	5.7	5.55	4.68	2.5	3.87	3.63	3.96	1.27	2.63	3.19
RS	NA	NA	NA	NA	NA	NA	NA	1	1.38	1	6.02	8.09	6.62	6.69	6.12	6.62	4.75	5.62	5.3	4.72	4.33	3.45
SE	1.09	1.05	1	1.03	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SI	2.82	2.64	2.53	2.62	2.72	2.37	2.4	2.54	2.89	2.01	2.24	2.45	2	1.84	1.76	1.97	1.32	1.49	1.37	1.52	1.39	1
SK	2.22	2.51	2.45	2.71	2.84	2.48	2.61	2.75	3.67	3.09	3.62	3.99	3.45	3.24	2.95	2.8	2.61	2.65	2.51	2.38	2.48	1.83
TR	2.89	3.02	3.27	3.54	3.69	3.29	3.55	3.24	3.16	2.97	3.53	4.26	4.08	3.79	4.22	3.56	2.99	3.37	2.85	2.88	2.82	1.97
UA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	1	1	1	1	1	1	1	1	NA	NA
UK	1	1	1	1	1	1.15	1.83	1.69	1.51	1.31	1.15	1.25	1.33	1.27	1.12	1.01	1	1	1	1.03	1	1

Note:
Output oriented model with constant returns to scale. Inputs are: rail_length, rolling_stock_pocketbook, staff_uic. Outputs are: rail_pkm, rail_tkm.

5.2 Second-Stage: Efficiency and Government Grants

This chapter estimates the model (see Equation 1) using the approach proposed by Simar and Wilson (2007) across four different specifications. All models are estimated in logarithms, with specifications (1) and (2) including fixed effects for each country and year.

To assess the impact of different funding structures - i.e., upstream vs. downstream grants - on efficiency, two new variables are constructed: total operating grants ($infra\_rev\_grant\_pps + psc\_pps$) and the ratio of upstream to downstream grants, $\left(\frac{infra\_rev\_grant}{psc}\right)$. The latter variable examines whether countries focusing more heavily on either upstream or downstream grants achieve higher performance scores. A higher ratio indicates that a greater share of funding comes from infrastructure revenue grants (e.g., subsidies for maintaining tracks) rather than public service obligations (compensation for running trains).

Specifications (1) and (3) examine grants per passenger-tonne-kilometer, while Specifications (2) and (4) focus on grants per capita. Both approaches have their advantages and disadvantages. Specifications (1) and (3) offer a performance-based perspective by linking funding to transport demand. However, it is highly sensitive to external shocks, such as the COVID-19 pandemic, which caused significant reductions in transport activity and potentially biased the measure. In contrast, Specifications (2) and (4) provide a more stable and consistent metric that is less influenced by short-term fluctuations in transport volumes. However, to account for potential structural breaks caused by the COVID-19 pandemic, interaction terms are included.

Specifically, the ratio variable is interacted with a binary indicator for the pre-pandemic period (*covid0*) and the pandemic/post-pandemic period (*covid1*), resulting in the terms $\left(\frac{infra\_rev\_grant}{psc}\right)$ *:covid0* and $\left(\frac{infra\_rev\_grant}{psc}\right)$ *:covid1.* These interaction terms enable an investigation of whether the relationship between funding structures and efficiency differs across these two distinct periods.

As shown in Table 3, the ratio of operating grants has a statistically significant negative effect on the efficiency scores in the pre-pandemic period, suggesting that higher revenue contributions to the infrastructure manager relative to public service compensation are associated with greater efficiency (as lower efficiency scores indicate higher efficiency). The coefficient is -0.035 in specification (1) and -0.034 in specification (2). This implies that a one percent increase in the ratio of infrastructure grants to public service compensation corresponds to an efficiency improvement of 0.035 % and 0.034 %, respectively. Specifications (3) and (4), estimated without year- or country-specific fixed effects, yield slightly larger coefficients of -0.037 and -0.067, respectively.

In other words: countries that grant a relatively higher share of operating grants to the infrastructure manager reach higher efficiency scores, while countries that focus their operating contributions on the compensation of transport undertakings seem to be less efficient. This can be explained as follows: It seems like the positive effect of revenue contributions, which typically lower access charges and thus stimulate intermodal and intramodal competition (e.g, Álvarez-SanJaime et al., 2016; Arrigo & Di Foggia, 2013), outweighs possible inefficiencies due to distorted cost structures (e.g., Obeng

& Sakano, 2020; Oum & Yu, 1994; Pucher et al., 1983). In addition, as a major part of compensation for public services is used to cover access charges, paying grants directly to the infrastructure manager can avoid double marginalization problems (e.g., Gutiérrez-Hita et al., 2022) and thus allow for a more efficient use of resources. However, the marginal effect of shifting funding to infrastructure on efficiency is relatively small, as (high) access charges are not the sole factor influencing competition (Crozet & Chassagne, 2013).

In contrast, the direction of the impact reverses in the post-COVID period. The coefficients for the interaction term are positive and statistically significant in specifications (1) and (2), with values of 0.124 and 0.123, respectively. In specifications (3) and (4), the coefficients are 0.021 and 0.064, which are smaller and not statistically significant. These results suggest that in the post-pandemic period, a higher ratio of infrastructure revenue grants to public service compensation is associated with a decline in efficiency. One possible explanation is that the disruption caused by the pandemic may have shifted priorities or led to inefficiencies in the allocation and utilization of operating grants, diminishing the benefits of stimulating competition through lower access charges. Furthermore, the pandemic likely exacerbated challenges in cost recovery for transport undertakings, making public service compensation a more critical factor for sustaining operational efficiency.

The analysis also reveals that infrastructure investment grants - whether measured per capita or per passenger-tonne-kilometer (ptkm) - have a significant and variable impact on efficiency. When measured in passenger-

Table 3: Results of the truncated regression - output oriented model, constant returns to scale (yearly frontier)

	(1)	(2)	(3)	(4)
log(infra_rev_grant/psc):covid0	−0.035**	−0.034**	−0.037***	−0.067***
	(0.016)	(0.015)	(0.013)	(0.014)
log(infra_rev_grant/psc):covid1	0.124***	0.123***	0.021	0.064
	(0.039)	(0.038)	(0.049)	(0.040)
log(infra_inv_grant_pps_ptkm)	−0.038**		0.048**	
	(0.017)		(0.020)	
log(infra_inv_grant_pps_capita)		−0.042**		−0.004
		(0.017)		(0.021)
log(infra_rev_grant_pps_ptkm + psc_pps_ptkm)	−0.137*		−0.116*	
	(0.070)		(0.066)	
log(infra_rev_grant_pps_capita + psc_pps_capita)		−0.186**		−0.284***
		(0.078)		(0.057)
log(gdp_pps_capita)	−1.236***	−1.111***	−0.550***	−0.123
	(0.268)	(0.255)	(0.120)	(0.135)
log(pop_density)	1.297**	0.952	−0.019	0.029
	(0.593)	(0.591)	(0.049)	(0.045)
log(share_freight)	0.166	0.114	−0.392***	−0.198**
	(0.196)	(0.191)	(0.106)	(0.098)
(Intercept)	5.738*	7.877**	5.574***	2.525**
	(3.208)	(3.281)	(1.108)	(1.194)
/sigma	0.071***	0.071***	0.163***	0.147***
	(0.005)	(0.005)	(0.013)	(0.012)
N	87	87	87	87
FE: country	yes	yes	no	no
FE: year	yes	yes	no	no

Note:
The dependent variable is log(efficiency_scores_out_crs_yearly). Negative coefficients indicate a positive effect on efficiency. Bootstrapped standard errors (2000 replications) are reported in parentheses. Coefficients are significant at the * 10%, ** 5%, and *** 1% level.

tonne-kilometers, the coefficients are −0.038 in specification (1) and 0.048 in specification (3), indicating differing effects across model specifications. Similarly, when measured per capita, the coefficients are −0.042 in specification (2) and −0.004 in specification (4), with the latter being statistically insignificant. These mixed results suggest that while infrastructure investment grants can positively influence efficiency under certain measurement and specification conditions, their effect is not consistent and may depend on contextual factors. One potential explanation for this variability is the strong correlation between costs and grants, as highlighted by Laabsch and Sanner (2012), which could obscure or complicate the interpretation of the grants' true impact on efficiency. On the other hand, the results in specifications (3) and (4) could also be driven by unobserved heterogeneity, which might be absorbed by the year- or country-specific fixed effects included in specifications (1) and (2). The coefficients for the country and year fixed effects are provided in Table B.1.

Moreover, the results demonstrate that total operating contributions (the sum of infrastructure revenue grants and public service compensation), measured either per capita (*infra_inv_grant_pps_capita*) or per passenger-tonne-kilometer (*infra_inv_grant_pps_ptkm*), can improve efficiency. The coefficients are −0.137 in specification (1) and −0.116 in specification (3) when measured per passenger-tonne-kilometer, and −0.186 in specification (2) and −0.284 in specification (4) when measured per capita. These findings indicate that higher operating contributions, regardless of the measurement approach, are associated with increased efficiency.

Finally, the non-financial control variables, namely GDP per capita (*gdp_pps_capita*), population density (*pop_density*), and the share of freight transport (*share_freight*), exhibit notable effects on efficiency. However, the size and significance of these effects vary significantly across the models. GDP per capita positively impacts efficiency in specifications (1)–(4), with coefficients of −1.236, −0.111, −0.550, and −0.123, although the latter is not statistically significant. As in other studies (e.g., Lerida-Navarro et al. 2019), the results suggest that wealthier economies tend to operate more efficiently. This finding can be attributed to several factors: higher GDP per capita is often associated with better-developed infrastructure, which reduces costs and enhances the reliability of transport operations. Moreover, wealthier countries are more likely to invest in advanced technology and innovation, fostering greater efficiency in transport systems. The availability of financial resources also supports higher operational standards and improved system management.

Population density has a negative effect on efficiency, as indicated by the coefficients in specifications (1), (2), and (4). In specification (3), population density shows a positive effect on efficiency. However, the coefficients are only statistically significant in specification (1). The negative relationship between population density and efficiency may reflect the challenges associated with densely populated areas. Higher population density often correlates with increased infrastructure congestion, leading to delays, overcrowding, and inefficient allocation of resources. These issues arise because the capacity of transport infrastructure may not keep pace with the high and diverse demand. Furthermore, the complexity of transportation networks in densely populated areas - encompassing urban, suburban, and intercity services - requires sophisticated coordination and incurs higher operational costs. While densely populated areas may benefit from economies of scale in transport services, these gains are often outweighed by the administrative and operational burdens tied to congestion and network complexity, ultimately reducing efficiency.

The influence of differences in the mix of freight and passenger transport on each country's network is not clear. In specifications (1) and (2), a higher share of freight transport is associated with a loss of efficiency. However, the coefficients are not statistically significant. In contrast, specifications (3) and (4), which were estimated without year- and/or country-specific fixed effects, suggest that a higher share of freight transport positively affects efficiency. This aligns with findings in the literature, which highlight that freight transport can have both positive and negative effects on efficiency. Fitzová and Nash (2024) argue that in countries with high freight volumes (e.g., bulk commodities), freight transport may benefit from economies of scale and standardized operations. However, a high share of wagonload traffic can reduce efficiency. Similarly, Lerida-Navarro et al. (2019) find that countries with a high modal share of rail freight transport tend to be more efficient.

It is important to note that these results may be influenced by the sample selection used in the second-stage regression. While the first-stage analysis includes a broad sample of countries, only eight countries are included in the second-stage regression due to data availability constraints. This limited sample may not fully represent the diversity of infrastructure quality, regulatory environments, and service mixes observed in the broader dataset. Consequently, the relationships identified in the second-stage regression should be interpreted with caution, as they may reflect the specific characteristics of the smaller sample rather than generalizable trends across all countries.

6 Conclusion

This study examines the effects of government funding structures on the efficiency of European railways, focusing on the allocation of grants to upstream infrastructure managers versus downstream service providers. Using a two-stage procedure involving Data Envelopment Analysis (DEA), I calculate efficiency scores and investigate their relationship with funding structures through a truncated regression model. The findings provide key insights into the implications of railway financing choices.

The results indicate that railways achieve higher efficiency when a larger share of government funds is allocated to upstream infrastructure managers. Specifically, a higher ratio of infrastructure revenue grants relative to public service compensation correlates with increased efficiency. This is attributed to lower access charges for downstream operators, which foster competition and enhance network utilization. Conversely, a reliance on downstream grants appears less effective, potentially due to inefficiencies linked to duplicated subsidies or insufficient incentives for cost containment.

Infrastructure investment grants exhibit mixed effects on efficiency. While these grants play a essential role in expanding capacity, improving network quality, and maintaining reliable operations, their impact varies depending on the model specification. In contrast, total operating contributions (the combined funding for infrastructure and public services) consistently demonstrate positive effects on efficiency.

In addition, non-financial variables such as GDP per capita and population density show varied effects on efficiency, underscoring the complexity of railway performance drivers. Wealthier countries benefit from advanced infrastructure and innovation, while densely populated areas face challenges from congestion and complex network coordination.

These findings underscore the importance of balanced, upstream-oriented funding structures to enhance railway efficiency. Policymakers should aim to optimize the allocation of public funds by emphasizing infrastructure support and fostering competitive, efficient rail networks aligned with broader environmental and economic goals.

References

Álvarez-SanJaime, Ó., Cantos-Sanchez, P., Moner-Colonques, R., & Sempere-Monerris, J. J. (2016). Rail access charges and internal competition in high speed trains. *Transport Policy*, *49*, 184–195.

Álvarez-SanJaime, Ó., Cantos-Sánchez, P., Moner-Colonques, R., & Sempere-Monerris, J. J. (2024). Efficiency versus market power in the rail industry. *Research in Transportation Economics*, *108*, 101477. https://doi.org/https://doi.org/10.1016/j.retrec.2024.101477

Arrigo, U., & Di Foggia, G. (2013). Competition and pricing of essential inputs: The case of access charges for the use of the italian rail infrastructure. *UTMS Journal of Economics*, *4*(3), 295–308.

Beck, A., Bente, H., & Schilling, M. (2013). Railway Efficiency - An Overview and a Look at Opportunities for Improvement. *International Transport Forum, Discussion Paper*, (2013-12).

Bernhofen, D. M. (1997). Strategic Trade Policy in a Vertically Related Industry. *Review of International Economics*, *5*(3), 429–433. https://doi.org/10.1111/1467-9396.00067

Boston Consulting Group. (2012). The 2012 European Railway Performance Index: Understanding what drives high performance by European railways: The Boston Consulting Group. Paris.

Cantos, P., Pastor, J. M., & Serrano, L. (1999). Productivity, efficiency and technical change in the european railways: A non-parametric approach. *Transportation*, *26*, 337–357.

Cantos, P., Pastor, J. M., & Serrano, L. (2002). Cost and Revenue Inefficiencies in the European Railways: WP-EC 2002-10. Instituto Valenciano de Investigaciones Económicas. Valencia.

Cantos, P., Pastor, J. M., & Serrano, L. (2012). Evaluating European railway deregulation using different approaches. *Transport Policy*, *24*, 67–72. https://doi.org/10.1016/j.tranpol.2012.07.008

Catalano, G., Daraio, C., Diana, M., Gregori, M., & Matteucci, G. (2019). Efficiency, Effectiveness, and Impacts Assessment in the Rail Transport Sector: A State-of-the-Art Critical Analysis of Current Research.

International Transactions in Operational Research, *26*(1), 5–40. https://doi.org/10.1111/itor.12551

Cavaignac, L., & Petiot, R. (2017). A quarter century of data envelopment analysis applied to the transport sector: A bibliometric analysis. *Socio-Economic Planning Sciences*, *57*, 84–96.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/https://doi.org/10.1016/0377-2217(78)90138-8

Coelli, T., & Perelman, S. (1996). Efficiency Measurement, Multiple-Output Technologies and Distance Functions: With Application to European Railways: CREPP 96/05 Draft. Centre de Recherche en Economie Publique et en Economie de la Population. Liège, Belgium.

Coelli, T., & Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: With application to European railways. *European Journal of Operational Research*, *117*(2), 326–339. https://doi.org/10.1016/S0377-2217(98)00271-9

Coelli, T., & Perelman, S. (2000). Technical efficiency of European railways: a distance function approach. *Applied Economics*, *32*(15), 1967–1976. https://doi.org/10.1080/00036840050155896

Coelli, T., Rao, P., O'Donnell, C., & Battese, G. (2005). *An introduction to efficiency and productivity analysis* (2nd ed). Springer.

Crozet, Y., & Chassagne, F. (2013). Rail access charges in france: Beyond the opposition between competition and financing. *Research in Transportation Economics*, *39*(1), 247–254.

European Commission and Directorate-General for Mobility and Transport. (2024). *Eu transport in figures – statistical pocketbook 2024*. Publications Office of the European Union. https://doi.org/doi/10.2832/16593

Färe, R., Grosskopf, S., & Margaritis, D. (2011). Malmquist productivity indexes and dea. *Handbook on data envelopment analysis*, 127–149.

Fiorentino, E., Karmann, A., & Koetter, M. (2006). The cost efficiency of German banks: a comparison of SFA and DEA: Deutsche Bundesbank

Discussion Paper, Series 2: Banking and Financial Studies, No 10/2006. Deutsche Bundesbank, Frankfurt.

Fischer, C., Greaker, M., & Rosendahl, K. E. (2014). Robust Policies against Emission Leakage: The Case for Upstream Subsidies: CESifo Working Paper no. 4742.

Fitzová, H. (2022). Has european rail policy improved the efficiency of european railways? *Journal of Transport Economics and Policy (JTEP)*, *56*(4), 445–463.

Fitzová, H., & Nash, C. (2024). Why are the railways of eastern europe less efficient than those of the west? *Transport Economics and Management*, *2*, 263–274. https://doi.org/https://doi.org/10.1016/j.team.2024.09.004

Friederiszick, H. W., Röller, L.-H., & Schulz, C. C. (2003). Evaluation of the effectiveness of state aid as a policy instrument: The Railway Sector. *Draft Version, 24.10.2003.*

Götz, G., & Schäfer, J. T. (2020). Financing railways. In *Handbook on railway regulation* (pp. 304–321). Edward Elgar Publishing.

Growitsch, C., & Wetzel, H. (2007). Testing for Economies of Scope in European Railways: An Efficiency Analysis: Working Paper Series in Economics No 72, University of Lüneburg, Germany.

Gutiérrez-Hita, C., de la Cruz, O., & Ramos-Melero, R. (2022). Infrastructure access charges, service differentiation, and strategic competition in the eu railway passenger market. *Transportation Research Part B: Methodological*, *164*, 87–104.

Hamilton, S. F., & Requate, T. (2004). Vertical structure and strategic environmental trade policy. *Journal of Environmental Economics and Management*, *47*(2), 260–269. https://doi.org/10.1016/S0095-0696(03)00080-9

Hokari, H., Hwang, H., & Ohta, H. (2003). Import tariff upstream versus export tax downstream: Are they welfare equivalent? *The Journal of International Trade & Economic Development*, *12*(3), 247–256. https://doi.org/10.1080/0963819032000132085

Holmgren, J. (2013). The efficiency of public transport operations – An evaluation using stochastic frontier analysis. *Research in Transportation Economics*, *39*(1), 50–57. https://doi.org/10.1016/j.retrec.2012.05.023

Kleinová, E. (2016). Does liberalization of the railway industry lead to higher technical effectiveness? *Journal of Rail Transport Planning & Management*, *6*(1), 67–76. https://doi.org/https://doi.org/10.1016/j.jrtpm.2016.04.002

Kyriacou, A. P., Muinelo-Gallo, L., & Roca-Sagalés, O. (2019). The efficiency of transport infrastructure investment and the role of government quality: An empirical analysis. *Transport Policy*, *74*, 93–102. https://doi.org/https://doi.org/10.1016/j.tranpol.2018.11.017

Laabsch, C., & Sanner, H. (2012). The impact of vertical separation on the success of the railways. *Intereconomics*, *47*(2), 120–128.

Lerida-Navarro, C., Nombela, G., & Tranchez-Martin, J. M. (2019). European railways: Liberalization and productive efficiency. *Transport Policy*, *83*, 57–67.

Lim, D.-J., Kim, M.-S., & Lee, K.-W. (2022). A revised dynamic data envelopment analysis model with budget constraints. *International Transactions in Operational Research*, *29*(2), 1012–1024.

Merkert, R., Smith, A. S., & Nash, C. A. (2010). Benchmarking of train operating firms – a transaction cost efficiency analysis. *Transportation Planning and Technology*, *33*(1), 35–53. https://doi.org/10.1080/03081060903429330

NERA. (2004). Study of the Financing of and Public Budget Contributions to Railways: A Final Report for European Commission, DG TREN. NERA National Economic Research Associates, London.

Niu, Y., Li, X., Zhang, J., Deng, X., & Chang, Y. (2023). Efficiency of railway transport: A comparative analysis for 16 countries. *Transport Policy*, *141*, 42–53. https://doi.org/https://doi.org/10.1016/j.tranpol.2023.07.007

Obeng, K., & Sakano, R. (2020). Effects of government regulations and input subsidies on cost efficiency: A decomposition approach. *Transport Policy*, *91*, 95–107.

Oum, T. H., Waters II, W. G., & Yu, C. (1999). A Survey of Productivity and Efficiency Measurement in Rail Transport. *Journal of Transport Economics and Policy*, (Volume 33, Part 1), 9–42.

Oum, T. H., & Yu, C. (1994). Economic Efficiency of railways and Implications for Public Policy: A comparative study of the OECD countries. *Journal of Transport Economics and Policy*, 121–138.

Pucher, J., Markstedt, A., & Hirschman, I. (1983). Impacts of subsidies on the costs of urban public transport. *Journal of Transport economics and policy*, 155–176.

Schäfer, J. T., & Götz, G. (2017). Public Budget Contributions to the European Rail Sector. *Review of Network Economics*, *16*(2), 89–123. https://doi.org/10.1515/rne-2017-0044

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), 31–64.

Simar, L., & Wilson, P. W. (2011). Two-stage dea: Caveat emptor. *Journal of Productivity Analysis*, *36*(2), 205.

Tretheway, M. W., Waters II, W. G., & Fok, A. K. (1997). The Total Factor Productivity of the Canadian Railways, 1956-91. *Journal of Transport Economics and Policy*, 93–113.

World Bank. (2011). Railway Reform: Toolkit for Improving Rail Sector Performance: The International Bank for Reconstruction and Development/The World Bank, Washington DC.

Appendix

A Descriptive Statistics

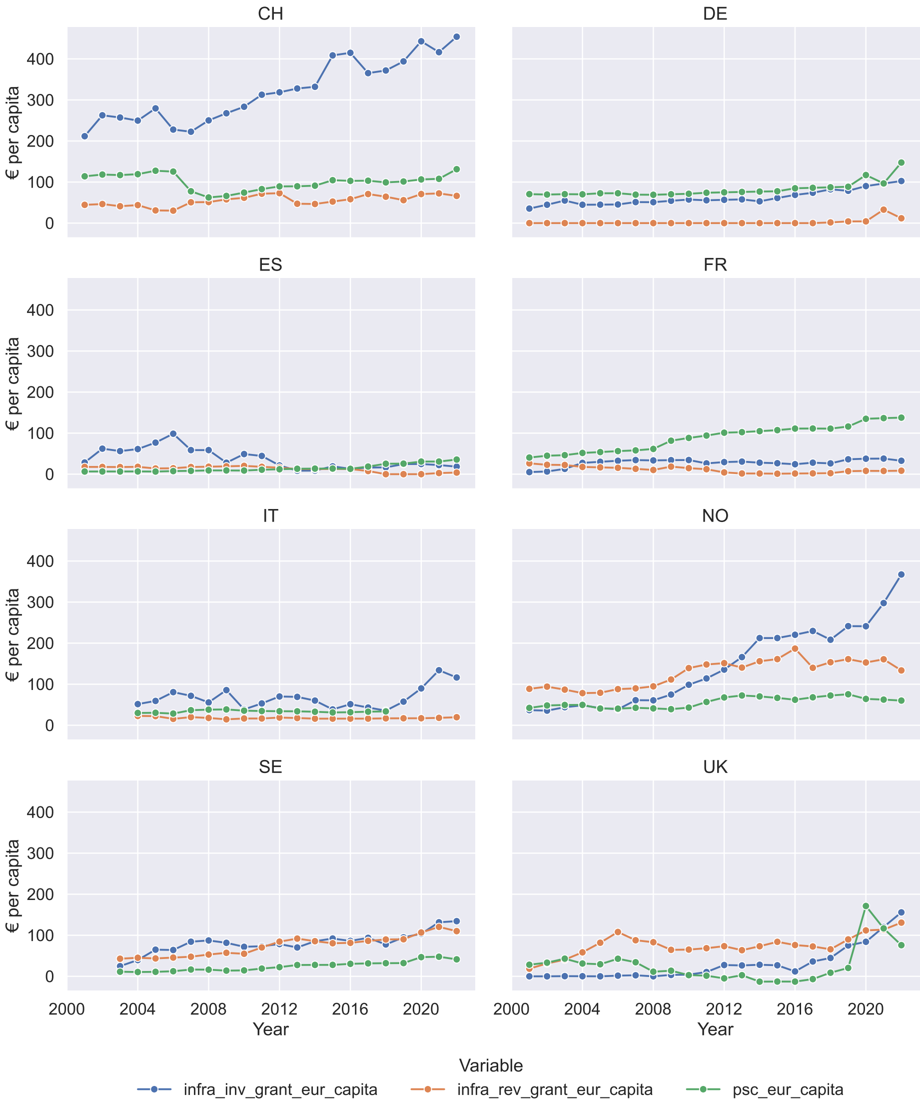


Figure A.1: Grants for infrastructure investment, the operation of the infrastructure, and the provision of public services in € per capita.

42

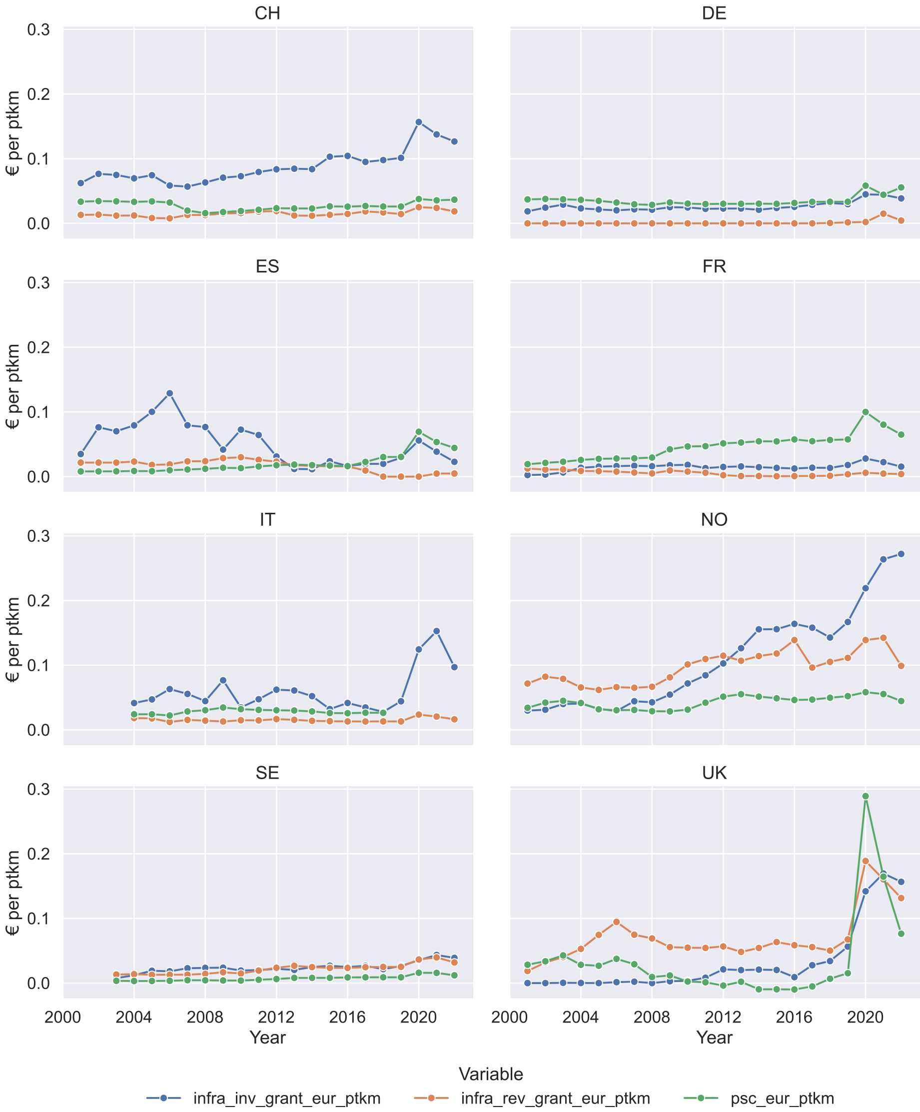


Figure A.2: Grants for infrastructure investment, the operation of the infrastructure and for the provision of public services in € per passenger-tonne-kilometer

B Robustness Checks

Table B.1: Results of the truncated regression - output oriented model, constant returns to scale (yearly frontier)

	(1)	(2)	(3)	(4)
log(infra_rev_grant/psc):covid0	−0.035**	−0.034**	−0.037***	−0.067***
	(0.016)	(0.015)	(0.013)	(0.014)
log(infra_rev_grant/psc):covid1	0.124***	0.123***	0.021	0.064
	(0.039)	(0.038)	(0.049)	(0.040)
log(infra_inv_grant_pps_ptkm)	−0.038**		0.048**	
	(0.017)		(0.020)	
log(infra_inv_grant_pps_capita)		−0.042**		−0.004
		(0.017)		(0.021)
log(infra_rev_grant_pps_ptkm + psc_pps_ptkm)	−0.137*		−0.116*	
	(0.070)		(0.066)	
log(infra_rev_grant_pps_capita + psc_pps_capita)		−0.186**		−0.284***
		(0.078)		(0.057)
countryDE	−0.214	−0.192		
	(0.147)	(0.140)		
countryES	0.858**	0.289		
	(0.430)	(0.471)		
countryFR	0.603*	0.284		
	(0.359)	(0.377)		
countryIT	−0.227*	−0.420***		
	(0.127)	(0.138)		
countryNO	3.539**	2.475*		
	(1.495)	(1.487)		
countrySE	1.968	1.210		
	(1.274)	(1.279)		
countryUK	−0.763***	−0.867***		
	(0.236)	(0.220)		
year2002	−0.001	0.001		
	(0.060)	(0.060)		
year2003	0.018	0.017		
	(0.061)	(0.060)		
year2004	0.119*	0.121*		
	(0.065)	(0.065)		
year2005	0.316***	0.312***		
	(0.073)	(0.070)		
year2006	0.350***	0.346***		
	(0.081)	(0.077)		
year2007	0.514***	0.508***		
	(0.089)	(0.087)		
year2008	0.405***	0.399***		
	(0.098)	(0.092)		
year2009	0.427***	0.409***		
	(0.107)	(0.103)		

year2010	0.424***	0.405***		
	(0.109)	(0.104)		
year2011	0.485***	0.475***		
	(0.108)	(0.103)		
year2012	0.554***	0.543***		
	(0.108)	(0.107)		
year2013	0.556***	0.545***		
	(0.113)	(0.109)		
year2014	0.525***	0.519***		
	(0.116)	(0.114)		
year2015	0.396***	0.394***		
	(0.121)	(0.114)		
year2016	0.399***	0.396***		
	(0.122)	(0.118)		
year2017	0.379***	0.377***		
	(0.128)	(0.125)		
year2018	0.249*	0.241*		
	(0.135)	(0.129)		
year2019	0.272*	0.266*		
	(0.144)	(0.141)		
year2020	0.606***	0.565***		
	(0.204)	(0.188)		
year2021	0.416**	0.384**		
	(0.163)	(0.156)		
year2022	0.477***	0.477***		
	(0.181)	(0.176)		
log(gdp_pps_capita)	−1.236***	−1.111***	−0.550***	−0.123
	(0.268)	(0.255)	(0.120)	(0.135)
log(pop_density)	1.297**	0.952	−0.019	0.029
	(0.593)	(0.591)	(0.049)	(0.045)
log(share_freight)	0.166	0.114	−0.392***	−0.198**
	(0.196)	(0.191)	(0.106)	(0.098)
(Intercept)	5.738*	7.877**	5.574***	2.525**
	(3.208)	(3.281)	(1.108)	(1.194)
/sigma	0.071***	0.071***	0.163***	0.147***
	(0.005)	(0.005)	(0.013)	(0.012)
N	87	87	87	87
FE: country	yes	yes	no	no
FE: year	yes	yes	no	no

Note:
The dependent variable is log(efficiency_scores_out_crs_yearly). Negative coefficients indicate a positive effect on efficiency. Bootstrapped standard errors (2000 replications) are reported in parentheses. Coefficients are significant at the * 10%, ** 5%, and *** 1% level.

Table B.2: Efficiency scores (pooled frontier) - output oriented model, constant returns to scale

country	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
AL	4.15	5.07	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
AT	2.56	2.4	2.43	2.32	2.28	2.06	2.03	1.94	2.23	2.09	1.93	1.94	1.9	1.81	1.79	1.73	1.7	1.67	1.03	2.21	2.03	1
BA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	8.85	9.04	5.08	9.57	8.85	5.46	6.65	5.71	5.26
BE	2.29	2.26	2.27	2.07	2.16	2.06	1.95	1.84	2.04	1.93	NA	NA	1	1.41	1.37	1.37	1.54	1.53	1.38	1.79	1.86	1.63
BG	5.9	6.49	6.05	6.1	6.02	5.85	5.81	6.32	8.17	8.18	7.97	8.84	6.94	6.61	6.19	6.58	5.87	5.92	5.79	5.25	4.96	4.27
CH	2.02	1.97	1.93	1.86	1.75	1.68	1.66	1.65	1.66	1.58	1.55	1.58	1.55	1.5	1.48	1.47	1.5	1.51	1.46	2.09	1.94	1.58
CZ	4.46	4.85	4.83	4.93	4.87	4.57	4.45	4.62	5.2	4.86	4.63	4.49	4.42	4.2	3.98	3.8	3.6	3.4	3.36	3.92	3.59	3.32
DE	2.26	2.39	2.7	2.56	2.26	2.12	2	1.98	2.17	2.07	2.01	1.99	1.95	1.94	1.89	1.79	1.84	1.82	1.79	2.43	2.26	1.8
DK	3.07	3.1	2.24	2.19	2.15	2.21	2.17	2.17	2.36	2.24	2.13	1.99	1.97	2.15	2.14	2.09	2.01	2.01	1.99	2.64	2.72	1.79
EE	1.21	1.02	1.09	1	1	1.04	1.25	1.43	1.35	1.2	1.3	1.62	NA	NA	2.94	3.64	3.65	NA	NA	NA	NA	NA
ES	3.3	3.24	3.29	3.33	3.33	2.9	3.02	2.87	3.08	3.27	3.14	3.06	3.08	2.83	2.41	2.34	2.29	2.23	2.13	4.29	3.37	2.2
FI	2.9	2.88	2.73	2.68	2.68	2.4	2.46	2.38	2.77	2.53	2.57	2.48	2.36	2.28	2.34	2.15	1.93	1.79	1.89	1.98	1.87	2.12
FR	2.36	2.32	2.37	2.31	2.31	2.23	2.19	2.1	2.19	2.17	2.07	2.09	2.09	2.04	1.96	1.96	1.87	1.9	1.82	2.86	2.19	1.68
HR	8.62	8.36	7.77	7.59	6.64	5.58	5.01	5.13	5.73	6	6.38	6.87	7.38	6.81	6.55	6.7	4.47	4.17	5.08	3.4	3.35	2.8
HU	4.24	4.06	4.16	3.97	3.96	3.95	4.03	3.68	4.5	4.3	4.24	4.14	3.88	3.66	3.69	3.61	3.06	3.46	3.44	3.61	3.62	3.29
IE	6.9	6.55	6.27	5.35	4.99	4.92	4.22	4.47	3.82	4.45	4.16	3.67	3.67	3.33	3.01	NA	2.72	2.53	2.44	7.18	6.84	NA
IT	2.01	2.04	2.07	2.01	2.01	2	2	2.05	2.23	2.25	2.25	2.23	2.18	2.11	2.02	1.99	1.95	1.87	1.84	3.42	2.77	2.06
LT	2.54	2.12	1.82	1.79	1.66	1.59	1.38	1.36	1.67	1.46	1.31	1.4	1.5	1.39	1.4	1.4	NA	1	1.11	1.11	1.15	NA
LU	3.31	3.9	4.05	3.98	4.57	4.08	3.52	4.35	4.85	4.15	4.27	4.3	NA	NA	NA	4.06	3.8	3.73	3.7	5.93	5.28	4.37
LV	1.65	1.51	1.26	1.28	1.2	1.38	1.26	1.16	1.17	1.24	1	1	1.11	1.13	1.16	1.34	1.38	1.08	1.21	1.99	2.01	NA
MD	NA	NA	NA	NA	NA	NA	NA	NA	9.26	9.71	8.74	10.26	8.73	9.54	11.89	14.89	12.59	12.25	12.22	NA	NA	8.83
ME	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	6.59	8.78	8.07	9.19	9.05	NA
MK	7.43	10	8.9	11.61	8.78	7.62	6.07	6.17	8.32	7.58	7.92	7.19	7.28	NA	NA	15.95	13.5	11.8	10.41	10.38	9.34	9.82
NL	1.29	1.3	1.32	1.13	1.08	1	1	1.09	1.11	1.14	1.09	1.04	1	1	1.07	1.09	1.07	1.03	1.01	1.8	1.57	1.12
NO	4.22	2.53	1.55	1.29	1.29	1.07	1.01	1	1	1	1.47	1.14	1.27	1.65	1.27	1.23	1.22	NA	NA	NA	NA	NA
PL	3.31	3.41	3.27	3.16	3.27	3.08	2.97	3.04	3.51	3.27	3.03	3.21	3.04	3.06	2.48	2.42	2.26	2.27	2.4	2.88	2.58	2.17
PT	4.17	4.16	4.33	3.89	3.62	3.4	3.17	3.13	2.08	3.29	3.15	2.94	3.03	2.85	2.11	2.53	2.47	2.42	2.34	3.56	3.72	2.75
RO	4.46	5.07	5.06	4.58	4.77	4.81	4.97	5.22	6.59	6.42	5.81	6.28	6.35	6.26	5.51	4.08	5.31	4.59	4.77	3.72	4.96	4.81
RS	NA	NA	NA	NA	NA	NA	NA	3.87	5.46	4.46	7.67	9.72	8.82	8.96	8.93	8.02	6.47	6.4	6.56	6.23	5.33	5.82
SE	2.15	2.13	2.01	1.95	1.86	1.79	1.62	1.74	1.84	1.39	1.4	1	1.03	1.02	1	1	1.01	1	1	1.05	1	1.08
SI	3.98	3.69	3.63	3.6	3.48	3.37	3.12	3.16	3.75	2.97	2.98	3.17	2.85	2.66	2.56	2.41	2.02	2.01	1.94	2.13	2	1.95
SK	3.32	3.47	3.64	3.75	3.8	3.6	3.68	3.71	4.45	3.98	3.98	4.09	3.75	3.58	3.39	3.38	3.25	3.23	3.22	4.38	3.93	3.49
TR	6.03	6.1	5.01	4.78	4.88	4.6	4.49	4.27	4.42	4.06	4.24	4.55	4.67	4.36	4.84	4.44	3.73	3.73	3.23	3.29	3.36	2.86
UA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	1.03	1.11	1.18	1	1.1	1.13	1.15	1.29	NA	NA
UK	1.48	1.41	1.13	1	1.3	1.36	1.9	1.88	1.85	1.76	1.43	1.33	1.34	1.27	1.28	1.28	1.27	1.26	1.22	2.92	2.53	1.78

Note:
Output oriented model with constant returns to scale. Inputs are: rail_length, rolling_stock_pocketbook, staff_uic. Outputs are: rail_pkm, rail_tkm.

Table B.3: Results of the truncated regression - output oriented model, constant returns to scale (pooled frontier)

	(1)	(2)	(3)	(4)
log(infra_rev_grant/psc):covid0	−0.046*	−0.057**	−0.030**	−0.034**
	(0.024)	(0.023)	(0.013)	(0.014)
log(infra_rev_grant/psc):covid1	−0.093**	−0.099**	−0.128***	−0.153***
	(0.041)	(0.039)	(0.028)	(0.028)
log(infra_inv_grant_pps_ptkm)	0.077***		0.120***	
	(0.025)		(0.018)	
log(infra_inv_grant_pps_capita)		0.080***		0.090***
		(0.023)		(0.022)
log(infra_rev_grant_pps_ptkm + psc_pps_ptkm)	−0.148*		−0.004	
	(0.083)		(0.038)	
log(infra_rev_grant_pps_capita + psc_pps_capita)		−0.277***		−0.140**
		(0.090)		(0.055)
log(gdp_pps_capita)	−1.661***	−1.677***	−1.105***	−0.958***
	(0.359)	(0.341)	(0.089)	(0.141)
log(pop_density)	1.584	1.501	0.038	0.024
	(0.966)	(0.940)	(0.026)	(0.026)
log(share_freight)	0.114	0.164	0.034	−0.040
	(0.228)	(0.219)	(0.086)	(0.084)
(Intercept)	9.691*	11.474**	12.332***	10.617***
	(5.322)	(5.298)	(0.995)	(1.286)
/sigma	0.136***	0.134***	0.182***	0.190***
	(0.009)	(0.009)	(0.012)	(0.014)
N	122	122	122	122
FE: country	yes	yes	no	no
FE: year	yes	yes	no	no

Note:
The dependent variable is log(efficiency_scores_out_crs). Negative coefficients indicate a positive effect on efficiency. Bootstrapped standard errors (2000 replications) are reported in parentheses. Coefficients are significant at the * 10%, ** 5%, and *** 1% level.

Table B.4: Results of the truncated regression - output oriented model, variable returns to scale (pooled frontier)

	(1)	(2)	(3)	(4)
log(infra_rev_grant/psc):covid0	−0.038	−0.056**	0.047*	0.046
	(0.025)	(0.024)	(0.027)	(0.032)
log(infra_rev_grant/psc):covid1	−0.082**	−0.092**	−0.086*	−0.146***
	(0.040)	(0.038)	(0.052)	(0.054)
log(infra_inv_grant_pps_ptkm)	0.046*		0.237***	
	(0.024)		(0.042)	
log(infra_inv_grant_pps_capita)		0.055**		0.192***
		(0.024)		(0.053)
log(infra_rev_grant_pps_ptkm + psc_pps_ptkm)	−0.093		−0.001	
	(0.088)		(0.069)	
log(infra_rev_grant_pps_capita + psc_pps_capita)		−0.247***		−0.272**
		(0.093)		(0.111)
log(gdp_pps_capita)	−1.535***	−1.582***	−0.852***	−0.592**
	(0.343)	(0.314)	(0.165)	(0.269)
log(pop_density)	2.191**	2.218**	−0.042	−0.082
	(0.939)	(0.893)	(0.049)	(0.060)
log(share_freight)	0.058	0.132	−0.034	−0.241
	(0.227)	(0.222)	(0.166)	(0.181)
(Intercept)	5.153	6.582	10.077***	6.895***
	(5.213)	(4.897)	(1.851)	(2.421)
/sigma	0.124***	0.121***	0.258***	0.278***
	(0.009)	(0.009)	(0.026)	(0.028)
N	118	118	118	118
FE: country	yes	yes	no	no
FE: year	yes	yes	no	no

Note:
The dependent variable is log(efficiency_scores_out_vrs). Negative coefficients indicate a positive effect on efficiency. Bootstrapped standard errors (2000 replications) are reported in parentheses. Coefficients are significant at the * 10%, ** 5%, and *** 1% level.

2.2. Entry deterrence due to brand proliferation: Empirical evidence from the German interurban bus industry

Authors: Samuel de Haas, Daniel Herold and Jan Thomas Schäfer

Published in: *International Journal of Industrial Organization*, vol. 83, 2022.

Please note that Fig. 1. in the manuscript does not show market shares before and after the takeover but rather average prices of the treatment and control group pre-merger (see also Fig. A.1. in the appendix of the manuscript). The correct figure can be found on page 81 of this thesis (Figure 2.1). The same holds for Fig. 2. in the manuscript which does not show the monopolistic and deterrent number of outlets but rather average frequencies of the treatment and control group pre-merger (see also Fig. A.2. in the appendix of the manuscript). The correct figure can be found on page 81 of this thesis (Figure 2.2).

International Journal of Industrial Organization 83 (2022) 102844

Contents lists available at ScienceDirect

International Journal of Industrial Organization

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

Entry deterrence due to brand proliferation: Empirical evidence from the German interurban bus industry

Samuel de Haas*, Daniel Herold, Jan Thomas Schäfer

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

ARTICLE INFO

Article history:
Received 23 December 2020
Revised 11 February 2022
Accepted 10 April 2022
Available online 29 April 2022

JEL classification:
L11
L41
L92
K21
K23

Keywords:
Competition
Takeover
Interurban bus services
Brand proliferation
Entry deterrence

ABSTRACT

We empirically study a potential entry deterrence strategy by means of product proliferation (pre-emption) in the German interurban bus industry using route-level data covering about 6,000 routes. In 2016, the largest operator, Flixbus, acquired its major rival, Postbus. We find that Flixbus, on average, provided a lower number of trips per route and day (frequency) and slightly *decreased* prices after the acquisition. Our findings are consistent with the interpretation that Flixbus pursued a strategy of pre-emption prior to the acquisition: Flixbus offered a high number of bus rides to decrease residual demand. After the acquisition, when the threat of entry was eliminated, Flixbus decreased the supply of trips per route and day. We would like to thank two anonymous referees and Frank Verboven for helpful comments and suggestions. We would also like to thank Niklas Dürr, Georg Götz, Justus Haucap, Paul Heidhues, Sven Heim, Michael Hellwig, Henry Thecat, Peter Winker and the participants of the following conferences: EARIE Annual Conference 2018, VfS Annual Conference 2018, MaCCI Annual Conference 2019 and CRESSE Conference 2019 for fruitful discussions on earlier versions of this article.

1. Introduction

The concept of entry deterrence and entry barriers has been analyzed by economists at least since Bain (1949). More thorough theoretical analyses using game theoretic concepts were then established by, for instance, Spence (1977), Dixit (1980, 1982), Kreps and Wilson (1982), Schmalensee (1982) and Judd (1985). Although there now exists a large body of theoretical literature[2], empirical analyses are scarce in this area. Empirical evidence is provided for specific industries such as automobiles (Barroso and Giarratana, 2013; Moreno and Terwiesch, 2016), pharmaceuticals (Bergman and Rudholm, 2003; Ellison and Ellison, 2011; Regan, 2008; Morton, 2000), hospitals (Dafny, 2005), airlines (Goolsbee and Syverson, 2008), hard drives (King and Tucci, 2002), casinos (Cookson, 2017), telecommunication (Weiman and Levin, 1994) and titanium dioxide (Hall, 1990). Cross-industry and overview studies are, for instance, presented by Bronnenberg et al., 2009, Bunch and Smiley (1992) or Bayus and Putsis (1999).

In this paper, we try to enrich the empirical knowledge on entry deterrence upon investigating a takeover in the interurban bus industry in Germany. In particular, we investigate a product proliferation or pre-emption strategy in the sense

* Corresponding author.
E-mail address: samuel.de-haas@wirtschaft.uni-giessen.de (S. de Haas).

[2] See, e.g., Bunch and Smiley (1992) or Gilbert (1989) for an overview of the various strands of literature.

https://doi.org/10.1016/j.ijindorg.2022.102844

of Schmalensee (1978), which means that the incumbent renders market entry unprofitable by committing to offer a large variety of products (Wilson, 1992). We interpret different rides on a given route as different variants of a good, i.e., a trip from Munich to Hamburg at 8 a.m. is no perfect substitute to the same trip at 4 p.m.. This gives the incumbent firm the opportunity to deter entry by offering trips at multiple times throughout the day, i.e., offering trips with a high frequency.

Various European countries have recently opened up competition in the markets for interurban bus passenger transport. One example is Germany, where the industry was deregulated in 2013.[3] After a period of market consolidation, by the end of 2016 Flixbus had a market share of more than 70% and the company announced the takeover of Postbus. Although Postbus was the largest remaining competitor of Flixbus, the operator was not active on all routes served by Flixbus. The merger was not investigated by the German antitrust authority because the turnover thresholds were not met.[4] The date of the takeover was November 1, 2016. At the same time, BerlinLinienBus, a subsidiary of Deutsche Bahn, left the market due to low profitability. As a consequence, Flixbus' market share rose to over 90% in the aftermath of the takeover.

We study the strategic behaviour of Flixbus prior to and after the takeover of Postbus. Our data set contains route-level data on prices and on the number of trips per route and day (frequencies) for about 6,000 routes in Germany for a period between September and December 2016 and the same period in 2017.

We find that after the takeover of Postbus *ceteris paribus* average prices increased and the average number of trips per route and day decreased. This is, however, mainly driven by the fact that Postbus was a low-price carrier, i.e., the prices were significantly lower than average. Controlling for this effect by using a fixed-effects panel regression with dummy-variables for the operator Flixbus and the takeover allows us to isolate Flixbus' behavior before and after the takeover. We identify a mild decrease in Flixbus' average prices of 0.55% – 0.7% and a drastic, highly significant decrease of 8.3 – 9.3% in Flixbus' average daily supply of trips per route.

While these findings are inconsistent with standard merger analysis (e.g., Farrell and Shapiro 1990), based on which one would expect that Flixbus increases prices due to the gains in market power, they can be explained by pre-emption in the context of a differentiated goods (Salop) model (Schmalensee, 1978; Eaton and Lipsey, 1979). That is, prior to the takeover Flixbus used its first-mover advantage to render market entry unprofitable on routes not served by Postbus by offering a high number of trips per day. By using this strategy, Flixbus was able to charge relatively high prices on those routes because the consumers' inconvenience costs were low. Given that the interurban bus market is rather small, Postbus was thus not able to establish a sufficiently large customer base to run a profitable business and, at some point, left the market. Additionally, entry of an entirely new firm was unlikely at least in the short run due to the necessary investments in marketing (establishing a customer base), a fleet and the permissions to serve the respective routes. After the acquisition of Postbus it was thus no longer necessary to follow a strategy of pre-emption so that Flixbus reduced the supply of trips per route and day on those routes where Postbus was not active before the acquisition. We also identify a statistically significant decrease in average prices of less than 1%. Even though the magnitude of the effect on prices is relatively small, it could be explained by an attempt to dampen a decrease in demand associated with the reduction of trips per day. Given that Postbus' business was basically shut down after the takeover, it is unlikely that the effect can be explained by merger efficiencies.

Our contribution to the literature is two-fold. First, we provide empirical evidence for the strategic use of pre-emption in a sense of product proliferation as in Schmalensee (1978). Although there are some empirical studies analyzing capacity pre-emption (Dafny, 2005; Ellison and Ellison, 2011) the closest related empirical study is Goolsbee and Syverson (2008).[5] They empirically investigate the incumbent firms' reaction to a rise in the threat of entry in the passenger airline industry in the US. They find that when the probability of entry increases incumbents decrease prices, which leads to higher passenger traffic. However, Goolsbee and Syverson (2008) cannot differentiate between specific flights, which our data allows us to do. The case we investigate is particularly interesting because the incumbent firm did not increase prices after the threat of entry was eliminated. We find that the high price level during the phase of entry deterrence was a result of the preemptive product proliferation strategy and the resulting large number of trips (variety) offered by the incumbent. Second, in the context of merger analysis, our study provides an example where constant or even slightly decreasing price levels after takeovers or mergers may actually be a sign of market power, especially in differentiated goods markets such as the German interurban bus industry. Decreasing prices alone are thus not necessarily an indication of welfare gains of a merger.

This article is structured as follows. In Section 2, we provide a brief overview of the German interurban bus industry. In Section 3, we present a theoretical model that is used to derive predictions for our empirical analysis. We present the data, descriptive statistics and empirical findings in Section 4. Section 5 concludes. Robustness checks and proofs can be found in the Appendix.

[3] Deregulation progresses at different rates across the member states of the EU. For instance, France deregulated the market at the end of 2015. Augustin et al. (2014) give a general overview of the industry key figures for Germany and the United States, where long-distance bus travel has been deregulated since the 1980s. van de Velde (2009) provides an overview about the regulatory framework of several European countries such as Germany, the U.K., Sweden and others. There are also various studies available examining the effects of market deregulation in single countries, e.g., Beria et al. (2014) (Italy), White and Robbins (2012) (U.K.) and Aarhaug and Fearnley (2016) (Norway).

[4] See a press release of the German Federal Cartel Office from August 3, 2018; https://bit.ly/37K63CZ.

[5] There is also another recent working paper by Bokhari and Yan (2019), which provides evidence for entry deterrence strategies by means of product proliferation in the UK pharmaceuticals market.

2. The German interurban bus industry

Several authors have analyzed the developments in the German interurban bus market. Dürr and Hüschelrath (2015) analyze the industry in 2015 with regards to the number of operators and the routes they provide, prices, market concentration as well as potential abuses of market power. Dürr et al. (2016) conduct an ex-ante analysis regarding the merger between MeinFernbus and Flixbus in 2015 that predicts increasing average prices post-merger. Although our findings regarding the takeover of Postbus differs from their prediction about the merger they analyze, there is a general consensus about Flixbus' dominant position in the market.

Based on a descriptive analysis of intramodal and intermodal competition in passenger transport, Knorr and Lueg-Arndt (2016) state that welfare gains would arise from the deregulation of the German interurban bus industry due to lower prices for passenger transport on average. Böckers et al. (2015) predict that competition between interurban bus and railway passenger transport operators leads to decreasing prices for long-distance rail passenger transport in Germany, which indicates that there exists a strong intermodal competition between the two modes of transport. Beestermöller (2017) identifies that railway strikes in Germany had a persistent, positive effect on the ticket sales in the interurban bus industry. Bataille and Steinmetz (2013) show that competition with interurban bus operators may decrease the profitability of railway operators, even on connections where no bus travel is provided. This effect arises from network effects in railway passenger travel. Our findings indicate that prices in the interurban bus industry significantly decrease on routes where there are alternative rail connections available. We therefore provide empirical evidence for the prevalence of intermodal competition between railway and interurban bus passenger transport.

In the following, we provide background information on the interurban bus industry in Germany. We present the evolution of the industry starting with the opening of competition in the market in 2013 until 2016, when Flixbus acquired Postbus.

2.1. Period I: prior 2013

Prior to the liberalization in 2013, the German interurban bus market was characterized by a very limited route network. This was due to regulations that had been introduced to protect rail traffic from competition. Only international connections and connections from and to former West Berlin were permitted. In addition to these connections there was a small number of routes for which suppliers could receive approval, e.g., Airport Shuttles. BerlinLinienBus, a Deutsche Bahn subsidiary, and Deutsche Touring were the two largest companies on the market at the time before the liberalization. The supplier DeinBus started to offer bus services on a regular basis using a legal loophole in December 2009. The connections were organized as carpool rides, and have been approved by a court decision in 2011. In April 2012, MeinFernbus began to operate its first route after an official approval was granted (see for example Dürr and Hüschelrath (2017) for an general overview of the industry development).

2.2. Period II: 2013 - 2014

In 2013, competition was opened up in the industry. According to the amended version of the German Passenger Transport Act, all connections may be approved if the distance between two successive stops is at least 50 kilometers. An exemption for shorter connections might be granted where no regional train connections with a travel time of less than one hour are available. Such an exemption furthermore requires that there are insufficient local transport services available or that only a small loss in demand for railway passenger transport is expected.[6]

The market for interurban bus transportation has grown rapidly after the liberalization process started, and several companies have entered the market. The resulting high competitive pressure forced some suppliers to leave the market or to merge with larger competitors. In February 2013, Flixbus entered the Market. City2City, a subsidiary of British National Express, and ADAC Postbus, a joint venture between the automobile club ADAC and Deutsche Post, entered in April and November 2013, respectively. In contrast to Flixbus and MeinFernbus, whose main business model was to offer their services to consumers with lower income such as students, ADAC Postbus addressed more demanding customers at that time (see Bundesamt für Güterverkehr 2015, 2016 for a comprehensive overview of entries and exits).

In October 2014, City2City was the first large operator to leave the market. One month later, in November 2014, ADAC withdrew from the joint venture with Deutsche Post. Deutsche Post thus continued to offer their services under the brand name Postbus. Postbus' business model of offering a high quality transport service for relative high prices turned out not be profitable. The operator then started to act more aggressively by offering lower prices and by regularly granting rebates. Although the number of suppliers started to shrink, Megabus, which is affiliated with the British Stagecoach Group, entered the market in December 2014 (Bundesamt für Güterverkehr, 2015).

[6] According to §42a Personenbeförderungsgesetz every connection with at least two successive stops more than 50 km apart is not considered short-range public transportation anymore.

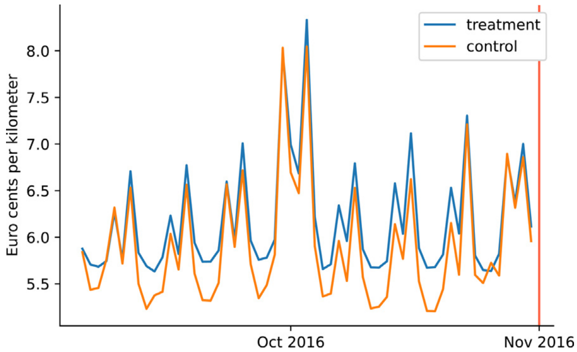


Fig. 1. Market shares before and after the takeover on November 1, 2016.

2.3. Period III: 2015 - fall 2016

At the beginning of 2015, the market structure changed significantly because of the merger of the two largest operators, Flixbus and MeinFernbus. The merged entity offered their services under the brand name Flixbus after a short transition period. In Summer 2016, Flixbus announced the takeover of Megabus. At this point the two main competitors of Flixbus were Postbus and BerlinLinienBus. However, BerlinLinienBus left the market in November 2016 due to low profitability. Finally, Postbus was taken over by Flixbus on November 1, 2016 (Bundesamt für Güterverkehr, 2016). This event is the object of our analysis.

Fig. 1 depicts the market shares on a monthly basis from September to December 2016. These figures are calculated using our data set which will be described in Section 4.1. The numbers indicate that Flixbus became a quasi-monopolist after the takeover. Only smaller competitors operating solely on a regional level (e.g., DeinBus) were also active in the industry. The licenses for Postbus' routes were acquired by Flixbus whereas a small number of BerlinLinienBus' routes were transferred to IC Bus, another subsidiary of Deutsche Bahn.

3. Theory

In this section, we present a simple model of the interurban bus industry prior to the takeover. The model provides the basis for our empirical analyses in Section 4. As described above, there was only a small number of competitors active in the market. Flixbus was the largest operator and was active on most routes in Germany whereas the smaller competitors' portfolios included only a limited number of connections. Given that our empirical analysis focuses on routes where Flixbus was the only active supplier, we consider Flixbus the incumbent firm on those routes.[7] The following model is therefore tailored to investigate the potential for entry deterrence on such a (representative) route where only the incumbent firm was active and where the threat of entry was immanent. We will show that entry deterrence can be achieved with high prices and a high number of bus rides on a given route (product proliferation) in equilibrium. In particular, in this equilibrium the monopolistic number of bus rides is lower than the deterrent number of bus rides so that the number of bus rides decreases after the takeover. Throughout the model, we assume that the incumbent commits to charge the price chosen on the first stage of the game. Without commitment on prices, the well-known concept of Limit Pricing occurs in equilibrium, where the incumbent firm chooses prices in a way that entry does not pay off (see, e.g., Milgrom and Roberts 1982), and which is thus closely related to our result.

We assume that bus rides are horizontally differentiated products. This is because consumers perceive alternative times of arrival as distinct variants of the transportation service. If a consumer has to meet an appointment at 3:00 pm, arriving at 2:30 pm might be a sweet spot for her. If the only available bus arrives at 2:55 pm she incurs inconvenience costs. Shy and Stenbacka (2006) impose a similar assumption on how consumers perceive shopping hours of supermarkets.

The incumbent is the first- and the entrant the second-mover as is standard in models of entry deterrence (see, e.g., Dixit 1982). That is, the incumbent first chooses prices and locations (e.g., departure times on the route). The entrant decides whether to enter on a subsequent stage. In case of entry, the entrant choose prices as well as locations.

As mentioned above, we assume that the incumbent commits not only to offer a certain variety but also to charge a certain price that cannot be changed afterwards. In that respect, our model differs from standard models of product differentiation where the first-mover only commits to offer a certain number of variants (or where variety is exogenously given) and where prices are determined in competition (Neven, 1987; Wilson, 1992; Salop, 1979b).[8] Our model thus allows for scenarios where the incumbent firms holds constant its market price (or, in other words, where prices are not influenced by potential entry) and decreases variety, i.e., the number of bus rides offered on a given route, when the threat of

[7] On routes where there was competition entry had already occurred (see Section 4.1).

[8] Price commitment is also analyzed, for instance, in Hagiu (2006) in the context of two-sided markets.

Table 1
Price determinants - fixed-effect panel regression results.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{Monday}	0.0524***	(0.0014)	0.0304***	(0.0013)
$D_{Wednesday}$	-0.0188***	(0.0011)	-0.0177***	(0.0011)
$D_{Thursday}$	0.0091***	(0.0011)	-0.0003	(0.0011)
D_{Friday}	0.1370***	(0.0018)	0.1316***	(0.0016)
$D_{Saturday}$	0.0519***	(0.0013)	0.0523***	(0.0012)
D_{Sunday}	0.1900***	(0.0022)	0.1853***	(0.0021)
D_{PH1}	0.0311***	(0.0008)	0.0292***	(0.0008)
D_{PH2}			0.1012***	(0.0013)
D_{PH3}			0.1235***	(0.0014)
D_{PB}	-0.3555***	(0.0059)	-0.3524***	(0.0060)
D_{TO}	-0.0032**	(0.0013)	0.0077***	(0.0015)
Constant	1.6648***	(0.0014)	1.6761***	(0.0015)
Observations	617,076		1,005,634	
R^2	0.6779		0.6655	
R^2 within	0.2856		0.2635	

The estimation is performed by using a GMM. Dependent variable is daily average price per kilometer and is expressed in logarithm. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

entry is eliminated. This closely resembles the idea of entry deterrence through brand proliferation as, e.g., described in Schmalensee (1978).

This assumption is motivated by the observations of market prices. Our data shows that Flxibus' prices were only marginally higher (approximately 1.7%) on routes where the firm was the sole provider in the pre-takeover period (see Table 5). Without price commitment one would expect to observe strong price differences on those routes where multiple providers are active, i.e., where entry has occurred. Notably, prices charged by Flixbus on a given route decrease more strongly when there are suitable train connections available to the consumers, which indicates that Flixbus' prices react more sensitively to prices of alternative train connections than to prices of alternative interurban bus connections. This provides evidence that, in order to establish a long-term alternative to train transport, the incumbent had an interest in offering stable prices. Refraining from price increases after becoming the dominant or even sole provider on a given route may prevent consumer search as, for instance, in Cabral and Fishman (2012), which could result in consumers switching to other modes of passenger transport such as train, car or plane.

If the incumbent firm could alter prices or the number of bus rides per connection after entry, any entrant would anticipate the incumbent's change in conduct as in Spence (1977). Provided that capacity is high enough – which is realistic because Flixbus relied on subcontractors[9] – the incumbent could technically behave like an unconstrained monopolist in this case. As will become clear below this is not consistent with our observations because Flixbus *reduced* the number of bus rides after the takeover.

The representative route is characterized by a circular city (Salop) model (Schmalensee, 1978; Eaton and Lipsey, 1979). There is a total number of L consumers uniformly distributed along a circle with unit circumference. For expositional convenience, and to maintain generality, we refer to Flixbus as the incumbent firm, indexed $i = 1$, and to Postbus as the potential entrant, indexed $i = 2$, throughout the following analysis. A single bus ride is referred to as an outlet in the context of the model. Prices are denoted p_i, $i \in \{1, 2\}$, and an outlet's location of firm i is denoted by l_i.[10]

The desired location of a given consumer is denoted l^*. The consumer incurs quadratic inconvenience costs per unit of distance $(l^* - l_i)^2$ as is the case in various Hotelling or Salop models (Bonanno, 1987; Neven, 1987).[11] Let u denote the consumer's gross valuation for the transport service. To simplify the analysis, we impose the assumption that demand is not "too small", i.e., that $u > 1$. This assumption rules out cases where the entrant has an incentive to locate its outlets midway between those of the incumbent, a fact that we will discuss below in more detail.

The consumer's utility function thus reads

$$U(l_i, l^*) = u - (l^* - l_i)^2 - p_i. \quad (1)$$

The marginal costs for transporting an additional passenger are zero. However, offering an additional bus ride entails fixed costs F (setup costs per outlet). Moreover, there are also (sunk) entry costs $S > 0$.[12] Those entry costs account, e.g., for

[9] see https://www.handelsblatt.com/english/companies/coach-travel-next-stop-monopoly/23568750.html
[10] It is sufficient to focus on two locations to derive the main results.
[11] Assuming linear transport costs makes the analysis much less tractable due to discontinuities in demand (d'Aspremont et al., 1979). Furthermore, one could multiply the transport costs with factor $c > 0$ as, e.g., in Bonanno (1987). We normalize $c = 1$ to simplify the analysis.

regulatory approval procedures. According to industry experts, the costs of obtaining approval to establish a new interurban bus connection in Germany ranges between a medium to a high six-figure sum.[13] Another example that illustrates the difficulties an entrant to the interurban bus industry faces is the British firm Megabus. The provider ordered buses worth 20 million pounds when entering the market in 2015. After just one year, Megabus left the market in 2016 (see Stagecoach Group plc annual report 2016, p. 16).

The utility obtained from consuming the outside good is normalized to zero.[14] The outside good is exogenously given and affects the market size of the interurban bus industry by capturing alternative transportation services. These alternatives are, first, going by car in which case utility is influenced by exogenous factors such as fuel prices. A second and third alternative would be traveling by plane and train, respectively. It is hard to imagine that prices and departure times of domestic flights respond to changes in the interurban bus industry. This is because the supply of domestic flights is limited to larger cities with an airport. Moreover domestic flights in Germany are typically targeting business customers due to their short traveling times and relatively high prices.[15] As we will show in Section 4.5, our empirical results suggest that the closest competitor to the interurban bus carriers is Deutsche Bahn. However, Deutsche Bahn is regulated and thus able to respond to the developments in the interurban bus market only to a limited extent.[16]

The timing of the game can be summarized as follows. On stage 1 the incumbent chooses a price p_1, the number of outlets, n, and the outlets' locations. On stage 2, the entrant decides whether to enter the market or not. In case of entry the entrant charges a price p_2 and chooses how many outlets to operate. If there is no entry, the entrant's profits are zero.

We now turn to our analysis. Consider first a situation where the incumbent behaves as an unconstrained monopolist. This constitutes the market environment without the threat of entry, i.e., after the takeover. Given the market is fully covered and the distance between each shop is equal the monopoly price satisfies

$$p(n) = u - \frac{1}{4n^2}. \tag{2}$$

Given (2), the monopolist's profit $\pi(p(n), n) = Lp(n) - Fn$ is maximized for

$$n_M \equiv \left(\frac{L}{2F}\right)^{\frac{1}{3}} \tag{3}$$

with the ensuing monopoly price (see, e.g., Eq. (5) in Bonanno 1987)

$$p_M \equiv u - \left(\frac{F^2}{16L^2}\right)^{\frac{1}{3}}. \tag{4}$$

Expressions (3) and (4) constitute the optimal number of outlets n_M and prices p_M from the viewpoint of an unconstrained monopolist.[17] If the incumbent chooses n_M and p_M when the entrant is active in the market, the latter may respond in a way that the incumbent may even be driven out of the market. Anticipating the entrant's response, it is reasonable to assume that the incumbent firm tries to deter entry.

In the literature, the canonical Hotelling location game where locations and prices are chosen on stages 1 and 2, respectively, two general cases are identified (see, e.g., Economides et al. 1986, Tirole 1988, Böckem 1994): (i) the entrant may prefer to operate its outlets close to those of the incumbent or (ii) entry may occur midway between the incumbent's outlets. The first and the second case can be interpreted as minimum and maximum product differentiation, respectively. Lemma 1 states the condition which determines whether, in case of entry, the entrant prefers to locate its outlets midway between or close to those operated by the incumbent. All proofs can be found in Appendix B.

Lemma 1. *The entrant prefers to locate its outlets midway between the incumbent's outlets if and only if* $p_1 \leq \frac{3}{4n^2}$.

The result can be interpreted as follows. A central driver of the entrant's incentive to induce minimum product differentiation is the business stealing effect. The magnitude of this effect decreases in the absolute distance between the firms' outlets.[18] That is, the business stealing effect becomes weaker the higher the perceived degree of product differentiation, which is captured by the consumers' transportation costs.

[12] As will be shown below, normalizing $S = 0$ implies that the incumbent realizes zero profits in case of entry deterrence. A similar result is shown in Salop (1979a). As pointed out by Geroski (1995) there has to be sunk entry costs for entry deterrence to be credible, which requires that it is profitable.

[13] See kcw-study "New Long-distance Coach Stations and Licensing Practices Opportunities for Municipalities" https://bit.ly/3np2Xtx (last accessed December 16, 2020).

[14] If the outside good yields positive utility, one can reformulate the problem by defining the willingness to pay for the good as the absolute utility from consumption minus utility obtained from consuming the outside good (Salop, 1979a).

[15] See, for instance, https://www.goeuro.com/travel/transport-price-index-2016 (last accessed on February 10, 2022).

[16] Deutsche Bahn's price schedules are typically changed on a yearly basis.

[17] Throughout the analysis we will treat the number of outlets n as a continuous variable. However, n is in fact an integer so that for any solution n^* one would have to compare profits for the floor and the ceiling of n^*. This can lead to slight deviations with respect to the result that the incumbent's profits have to be at or below entry costs S to deter entry (see below). However, similar to Salop (1979a), this does not qualitatively impact our results because neither does the incumbent incur losses nor is there enough free space for profitable market entry when the optimal, integer-valued number of outlets is $\lfloor n^* \rfloor$ or $\lceil n^* \rceil$.

[18] To see this, consider $l_1 = 0$ and any location $l_2 > 0$. The indifferent consumer is then located at $z = \frac{l_2}{2} + \frac{p_2 - p_1}{l_2}$. The term $\frac{p_2 - p_1}{l_2}$ captures the business stealing effect.

According to Lemma 1 there exists an upper bound $\frac{3}{4n^2}$, which ensures maximum product differentiation. This upper bound decreases in n. That is, it becomes less likely that the entrant locates as far away as possible from the incumbent because, when n increases, the perceived distance between the incumbent's variants becomes smaller and the incentive for business stealing becomes stronger.

It follows from Lemma 1 that cases where the entrant locates midway between the incumbents' outlets cannot occur when the incumbent covers the whole market. The price that ensures full market coverage is depicted in (2). Given $p(n)$ in (2) and according to Lemma 1 midway-entry requires $u \leq \frac{1}{n^2}$. Given that $n \geq 1$ is a necessary assumption for the existence of a market and given the assumption that demand is not too small ($u > 1$), it always holds that $u > \frac{1}{n^2}$. Thus, in case of entry the entrant will always locate its outlets close to those of the incumbent or, in other words, in an equilibrium with entry there will always be minimal product differentiation.

In the light of Lemma 1, increasing the number of variants n has the familiar effects inherent to models of entry deterrence with differentiated goods. On the one hand, the incumbent is able to charge higher prices by introducing new variants because consumers' transportation costs decrease. On the other hand, when prices are high the incentive for the entrant to be active in the market by introducing variants similar to those of the incumbent is relatively strong.

In order to examine the case of entry deterrence, the number of outlets and the price in case of entry deterrence are defined by n_D and p_D, respectively. As shown above, in cases where entry occurs close to the incumbent's chosen locations the entrant marginally undercuts the incumbent. The entrant will then serve all customers. Entry deterrence thus requires that the following condition holds.

$$L(p_D - \epsilon) - Fn_D - S \leq 0 \tag{5}$$

Condition (5) states that the entrant's profit from charging a price $p_D - \epsilon$, $\epsilon \to 0$, has to be weakly negative for entry to be deterred. By locating each outlet marginally close to the incumbent's outlets, the entrant's number of outlets is n_D. In addition to the setup costs F for each outlet, the entrant incurs entry costs S. It also follows from (5) that the incumbent's profit in case of entry deterrence, $Lp_D - Fn_D$, has to be approximately equal to or below S because otherwise entry occurs, which is consistent with the concept of "limit profits" in the context of Limit Pricing (Geroski, 1995, p. 427).

In order to ensure that condition (5) is satisfied, the incumbent can adjust two variables, i.e., the price p_D and the number of outlets n_D. This implies that the optimal incumbent's strategy of entry deterrence is not uniquely determined. It is thus possible to obtain results where prices decrease or increase after the takeover, as will be shown below. Consider first a situation where the incumbent charges the monopoly price p_M and accordingly operates n_D outlets in order to deter entry.

Entry deterrence by charging the monopoly price p_M occurs when the number of outlets n_D exceeds the number of outlets an unconstrained monopolist would choose, n_M.[19] Lemma 2 states a lower bound $\underline{n}$ on the number of outlets above which entry is deterred, i.e., for all $n_D \geq \underline{n}$, given the price is p_M.

Lemma 2. *The monopoly price p_M as defined in (4) occurs in an equilibrium with entry deterrence if and only if n_D exceeds*

$$\underline{n} = \frac{Lu - S}{F} - \left(\frac{L}{16F}\right)^{\frac{1}{3}}. \tag{6}$$

The lower bound of the number of outlets $\underline{n}$ as depicted in (6) decreases in the entry costs S. This implies that the higher entry costs S the lower the minimum number of outlets in order to deter entry, i.e., it becomes easier for the incumbent to defend its monopoly position. Given that an unconstrained monopolist's number of outlets n_M (see (3)) is independent of S, the result that the incumbent's number of outlets decreases after a takeover is *less* likely to occur when entry costs S are high. Formally, $\underline{n} \geq n_M$ follows if and only if

$$S \leq Lu - \frac{3}{2}\left(\frac{F^2 L}{2}\right)^{\frac{1}{3}}. \tag{7}$$

To illustrate our result, consider the following parameter values. Assume the daily market for an interurban bus service that connects two cities that are 800 km apart (for instance, Munich and Hamburg). One could assume that, per day, there may be a total number of $L = 1,000$ potential bus travelers, each one willing to pay at most $u = 50$ per bus ride. Suppose that the costs of providing a single bus on this route is $F = 500$. The ensuing monopoly price is $p_M = 49.75$, which yields a price per km of 6.22. This price (in € -cents) is in line with the empirical observations we will present in Section 4.2. Given these parameter values, the upper bound of entry costs S below which the result $\underline{n} \geq n_M$ occurs is 49,250 (see (7)). Noting that $p_M > \frac{3}{4n^2}$ (see Lemma 1), minimum product differentiation always occurs in the relevant range where $\underline{n} \geq n_M$. Fig. 2 shows the minimum number of outlets $\underline{n}$ as a function of S.

The solid line in Fig. 2 depicts the monopolistic number of outlets, n_M, which is one in our example. The dashed line depicts $\underline{n}$ as explained above. That is, for all S below the intersection point of the solid and the dashed line, it holds that $n_M \leq \underline{n}$ so that the number of outlets in case of deterrence ($n_D \geq \underline{n}$) would be (weakly) greater than the number of outlets

[19] Recall that with p_M and n_M the market was fully covered thus the market will also be fully covered for $p_D = p_M$ and $n_D \geq n_M$.

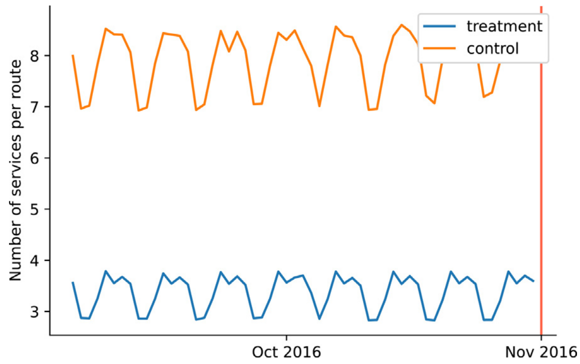


Fig. 2. Monopolistic and deterrent number of outlets.

chosen by a monopolist. In that range, variety would decrease after a takeover. The dotted line shows the minimum number of outlets $\underline{n}'$ such that the incumbent can charge a deterrent price of $p_D = p_M + 0.5$, i.e., a price which is roughly 1% above the monopoly price. This visualizes the incumbent's opportunity to deter entry for a range of different prices. One can see that entry can be deterred with a higher price than the monopoly price and that this is also consistent with a decrease in variety (i.e., a lower number of bus rides per day and route) when sunk costs are approximately below $S = 49,750$.[20]

At this point it becomes obvious that our model is also consistent with the concept of Limit Pricing, i.e., if the incumbent chose a price other than p_M in order to maintain a monopolist's number of variants n_M in case of entry deterrence. From (5), there exists an upper bound (limit price) for the deterrent price $p \leq \frac{S}{L} + \left(\frac{F^2}{2L^2}\right)^{\frac{1}{3}} \equiv \overline{p}$. Accordingly, an increase in prices after the takeover, i.e., $\overline{p} \leq p_M$, occurs if (7) holds. In our model, a price increase after the takeover is therefore equivalent to a decrease in the number of variants.

Assuming the incumbent holds constant the prices in the pre- and post-takeover-period, our theoretical model predicts that entry deterrence in the sense of product proliferation occurs when Flixbus offers a larger number of bus rides in the pre-takeover-phase compared to the post-takeover-period. Such a pattern is consistent with a drop in the supply of bus rides on a given route after the takeover.

4. Empirical analysis

In this section, we will present our empirical analysis. We first describe our data set and show descriptive statistics. In a second step, the main part of the empirical analysis will be presented, i.e., the effects of the takeover.

4.1. Data

The data set underlying our analyses was constructed based on busliniensuche.de, one of the leading online search engines for interurban bus travel in Germany. Our data set includes information on travel dates, providers, origins and destinations as well as on durations and prices of more than 8,000 routes in Germany.[21] After excluding routes that are offered less than once per week and provider, the data set contains 5,942 routes as observations. The search engine also provides data on train connections of Deutsche Bahn, which we use to examine how railway connections available to consumers as a substitute on a given bus route affect prices and frequencies in the interurban bus industry.

We consider three groups of providers, i.e., Flixbus, Postbus and Others. The latter group contains all operators except Flixbus and Postbus (e.g., BerlinLinienBus or Deinbus). The first part of the data set contains information on all providers, all available routes and each day from September 5 to December 11, 2016, i.e., in the year of the takeover (henceforth referred to as the takeover-period). To control for seasonal effects (excluding Christmas, which we exclude from our data set), which occur over the course of the year, the second part of the data set contains data on the period September 5 to December 11, 2017, i.e., one year after the merger (henceforth referred to as the post-takeover-period).

All information on transportation services were collected by requesting the transportation service five days prior to the respective day of departure.[22] We started collecting the data before BerlinLinienBus announced its market exit. It is thus

[20] The result can be derived as follows. Suppose that the deterrent price is $p_D = p_M + x$. To ensure deterrence, $L(p_M + x - \epsilon) - Fn_D - S \leq 0$ (see (5)), which yields $\underline{n}' \equiv \frac{L(u+x)-S}{F} - \left(\frac{L}{16F}\right)^{\frac{1}{3}}$. With the given parameter values and $x = 0.5$, the intersection point between $\underline{n}'$ and $n_M = 1$ is $S = 49,750$.

[21] Following Dürr et al. (2016), we define a route as each combination between two different stops on a connection (e.g. Hamburg to Munich) and count outward and inward trips as two separate routes. If one provider offers the same route more than once per day, we aggregate the data on a provider-level per day.

[22] Dürr and Hüschelrath (2015) use a similar approach. They do not find substantial variation in the data when the time spans between booking and the day of departure differ.

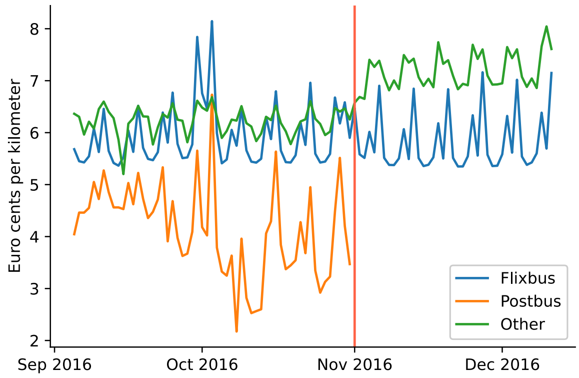


Fig. 3. Average prices by operator in € -cents per km.

not possible to distinguish between BerlinLinienBus and other providers contained in the group Others in our data set. This does not affect the economic interpretations of our results because, against the background of our theoretical model, another potential entrant left the market and we consider only routes where only Flixbus is active.

4.2. Descriptive statistics

The average price of an interurban bus service was € 15.66 prior to the takeover and the average price per km was € -cents 5.67. The evolution of the average prices is shown in Fig. 3. The vertical line represents the date of the takeover. In order to analyze the competitive environment before the takeover, each route is assigned to one of 6 subgroups as follows.

1. 3,707 routes provided only by Flixbus (on average € -cents 6.09/km)
2. 238 routes provided only by Postbus (€ -cents 4.46/km)
3. 444 routes provided by Flixbus and Postbus (€ -cents 4.96/km)
4. 638 routes provided by Flixbus and Others (€ -cents 6.33/km)
5. 121 routes provided by Postbus and Others (€ -cents 5.78/km)
6. 794 routes provided by Flixbus, Postbus and Others (€ -cents 5.07/km)

After the takeover, three types of routes remain. First, there are routes on which only Flixbus is active. These routes were the former subgroups 1, 2 and 3. From the subgroups 4, 5 and 6 emerge, second, a set of routes on which Flixbus and the group Others and, third, only the group Others are active after the takeover. Note that, regarding subgroup 2, only 5 routes served by Postbus before the takeover were adopted by Flixbus. The following routes (5,197 in total) were available after the takeover.

- 4340 routes provided only by Flixbus (on average € -cents 5.91/km).
- 782 routes provided by Flixbus and Others (€ -cents 6.07/km).
- 75 routes provided only by Others (€ -cents 14.16/km)[23].

Before turning to a more detailed analysis of the evolution of prices in the industry, it seems appropriate to first examine the overall development of the industry before and after the takeover. The following figure summarizes the industry-wide price effects of the takeover on an aggregate level. Fig. 3 shows the average prices of each operator in € -cents per km in the takeover-period. Postbus left the market on November 1, 2016. One can see that although the other operators increased their prices on average, it seems that the prices charged by Flixbus remained almost constant after the takeover.[24]

Average prices in the industry increased from approximately 5.67 € -cents before to 6.02 € -cents after the takeover. There are two reasons for this increase. First, Postbus and BerlinLinienBus were low-price carriers. If these suppliers leave the market *ceteris paribus* industry average prices will unambiguously increase. Second, Flixbus or the remaining suppliers (Others) may have increased prices as a consequence of the reduction in the number of operators in the market. In what follows, we isolate these effects and evaluate their impact on the industry after the takeover.

4.3. Prices

In this section, we examine the impact of the takeover on prices. Industry-wide prices are analyzed using a fixed-effects panel regression. The structural equation takes the following form:

$$\ln p_{i,j,t} = X_t'\beta_1 + \delta_1 D_{TO,t} + \delta_2 D_{PB,j} + \alpha_i + \epsilon_{i,j,t}. \quad (8)$$

[23] Average prices in this group are relatively high to cover the high average costs due to low demand for most of these, e.g., small sections of international connections.

[24] The increase in prices on October 3, 2016 and the previous weekend can be explained by public holidays.

Table 2
Price determinants group 1 - fixed-effect panel regression results.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{Monday}	0.0465***	(0.0013)	0.0192***	(0.0012)
$D_{Wednesday}$	-0.0172***	(0.0009)	-0.0176***	(0.0008)
$D_{Thursday}$	0.0118***	(0.0013)	-0.0037***	(0.0012)
D_{Friday}	0.1323***	(0.0022)	0.1230***	(0.0020)
$D_{Saturday}$	0.0440***	(0.0014)	0.0408***	(0.0014)
D_{Sunday}	0.1988***	(0.0030)	0.1905***	(0.0028)
D_{PH1}	0.0545***	(0.0009)	0.0535***	(0.0010)
D_{PH2}			0.0989***	(0.0016)
D_{PH3}			0.1260***	(0.0019)
D_{TO}	-0.0060***	(0.0010)	-0.0009	(0.0014)
Constant	1.6865***	(0.0012)	1.7087***	(0.0015)
Observations	291,532		492,075	
R^2	0.7684		0.7333	
R^2 within	0.2124		0.2108	

The estimation is performed by using a GMM. Dependent variable is daily average price per kilometer and is expressed in logarithm. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * p < 0.05, ** p < 0.01, *** p < 0.001.

In Eq. (8), the dependent variable is the logarithm of daily average prices per km. Prices are given for each route i and provider j on day t. The matrix X_t includes dummy variables for each day of the week (Dürr et al., 2016). Given that October 3, 2016 (and 2017), was a national public holiday in Germany, we also include dummy variables to control for potential effects on prices arising from that holiday. In 2017 there also was a national public holiday on October 31. For this purpose we specify 3 public holiday dummies, which take the value 1 in the week before and the week after the respective public holiday and 0 for the remaining observations.[25] The dummy variable $D_{TO,t}$ is 0 before and 1 after the takeover. The coefficient δ_1 therefore captures the effect of the takeover on prices of the remaining operators. We also include a dummy variable $D_{PB,j}$ that takes the value 1 if the corresponding provider is Postbus and the value 0 otherwise. The coefficient δ_2 thus estimates the (percentage) difference between prices of Postbus and the industry average. Time-invariant heterogeneity between different routes, e.g., sociodemographic characteristics of the cities, is absorbed by the route fixed-effect α_i. Table 1 presents the results of regression (8) for the takeover-period as well as for the post-takeover-period.[26]

We find a strong increase in average prices on weekends of roughly 14% (5.19%, 19.00%) on Fridays (Saturdays, Sundays) in 2016 and similar magnitudes in 2017. This is in line with Dürr et al. (2016) who find that demand for interurban bus services is higher on weekends. Over the course of the public holiday of October 3, 2016 we observe an average price increase of 3.11%. In 2017 this increase was even more pronounced: prices approximately increased by an additional 10.12% (12.35% for the public holiday on October 31) relative to normal weekdays.

We distinguish between a short-term and a long-term perspective. In both cases, the pre-takeover period is September 2016 – October 2016. From the short-term perspective, we investigate the effects of the takeover on frequencies only for the period November 2016 – December 2016. The long-term perspective, for which the post-takeover period is November 2016 – December 2016 and September 2017 – December 2017, reveals that the effects of the takeover were not just temporary effects that lasted for a few month. The focus lies on the periods November – December 2016 and September 2017 – December 2017 to ensure comparable seasonal effects.

The coefficient δ_1 is negative and close to zero (-0.32%), thus, there is little indication that the average prices of the remaining competitors increased in the takeover-period. Over a longer period of time, i.e., using data for the period November – December 2016 and September – December 2017 (long-term perspective), we identify a small increase in prices of about 0.77%. Moreover, as indicated above, there is also a direct effect on average prices because the low-price operator Postbus left the market. That Postbus was indeed a low-price operator is confirmed by the coefficient δ_2, which indicates that that operator charged prices approximately 35% below the average price level of the industry.

We now examine how the prices charged by Flixbus changed after the takeover. Especially the routes of group 1 are important for our analysis because these routes are served only by Flixbus before and after the takeover. However, there is no longer the threat of entry after the takeover. Changes in prices are analyzed using estimation Eq. (8) (without the dummy $D_{PB,j}$ because only data of group 1 are included). Table 2 captures the corresponding results for both the takeover-period and the post-takeover-period.

[25] D_{PH1} takes the value 1 for observations between 2016-09-26 and 2016-10-09, D_{PH2} takes the value 1 between 2017-09-25 and 2017-10-08 and D_{PH3} takes the value 1 between 2017-10-23 and 2017-11-05.

[26] Note that we run our estimations by means of a balanced panel. We only include routes that were available before and after the takeover. However, including also routes that were no longer served after the takeover leads to (qualitatively) the same results. A full list of controls included in the estimation can be found in Appendix A.

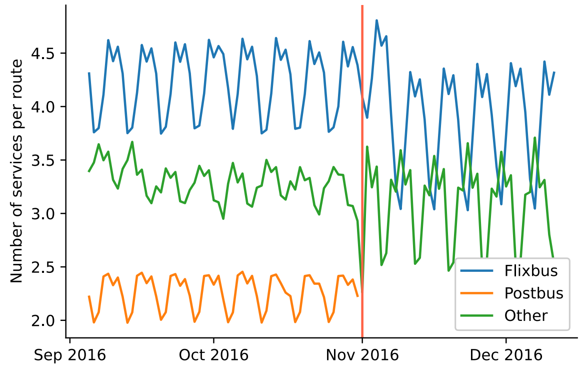


Fig. 4. Number of daily trips per route by group and operator.

One can see that the coefficients of the dummy variable D_{TO}, which correspond to δ_1 in (8), are small but significantly negative. The results indicate that, from a short-term perspective, Flixbus *decreased* its average prices after the takeover by 0.60% compared to the level of the pre-takeover period. Taking into account the period September – December 2017 (long-term perspective) the coefficient is not significantly different from zero. As discussed above, based on standard merger analysis one would expect that, in general, Flixbus would increase prices after the takeover.[27]

In general, decisions to take over a competitor or to exit a market are driven by expected profits and, therefore, by prices. To test whether the results of our estimations are affected by endogeneity, we apply the heteroscedacity based instrumental variable (IV) approach suggested by Lewbel (2012). There is no indication that our results might be affected by endogeneity because the results show only minor differences to the results presented in this section. Additionally, our results might be driven by, first, seasonality: On the one hand, an exogenous trend, e.g., decreasing demand might drive our results. Second, seasonal volatility (intra-year fluctuations) might also drive our results. To check for robustness of our analyses against the first hypothesis, we applied a difference-in-differences approach, where the routes of group 1 represent the treatment group and the routes of group 4 the control group. To account for a potential bias stemming from the exit of BerlinLinienBus, we keep only such routes in the control group that remain competitive and continue to be served by other operators. Furthermore, we estimated the above stated regression using only data from the months September and October (both, 2016 and 2017) and we could not find any signs of seasonal volatility. The corresponding results are presented in more detail in Appendix A and indicate that the results presented in this section are robust.

Our results indicate that due to the exit of Postbus and BerlinLinienBus the average price level in the industry increased after the takeover. However, we do not find evidence that Flixbus strongly increased its own prices after the takeover. Flixbus also did not increase its prices in the year after the takeover.

4.4. Frequencies

Our theoretical model predicts that the number of trips per route and day (frequencies) decreases after the takeover. The frequencies of each operator are examined in the following.

Fig. 4 displays the number of trips per route and day for each operator. The industry-wide average daily number of trips over all routes and providers was about 5.31 before and 4.23 after the takeover in the takeover-period, i.e., there is a decline in the supply of bus rides of about 20%. However, as is the case for prices in the industry, this observation can partly be explained by the discontinuation of services formerly offered by Postbus. However, we also observe that all remaining providers reduced the number of trips per route and day. Note that this is not what we would expect in a Cournot-market, where the output of the remaining firms would increase when a supplier leaves the market. To further examine this finding, we first perform a fixed-effects panel regression on the industry-level, i.e., containing data on all routes. The structural equation of our model takes the following form:

$$\ln q_{i,j,t} = X_t' \rho_1 + \lambda_1 D_{PB,t} + \lambda_2 D_{TO,t} + \alpha_i + \nu_{i,j,t} \tag{9}$$

In Eq. (9), the dependent variable is the logarithm of the number of trips provided by firm j on route i and day t. The matrix X_t includes dummy variables for each day of the week as well as dummy variables for the public holidays on October 3, 2016 (and 2017) as well as October 31, 2017. As in the analysis of prices the dummy variable $D_{PB,t}$ is 1 if the supplier of the respective route is Postbus and zero otherwise. The dummy variable D_{TO} takes the value 0 (1) before (after) the takeover. The coefficients λ_1 and λ_2 thus capture the difference in the daily supply of routes by offered Postbus compared

[27] Price decreases are possible when the takeover, or merger, creates synergies reducing the operator's marginal costs (Farrell and Shapiro, 1990). Given that Postbus' business was basically shut down it seems unlikely that such synergies resulted from the takeover.

Table 3
Frequencies - fixed-effect panel regression results.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{Monday}	0.1167***	(0.0042)	0.1259***	(0.0040)
$D_{Wednesday}$	-0.0136***	(0.0026)	-0.0205***	(0.0029)
$D_{Thursday}$	0.0991***	(0.0041)	0.1013***	(0.0039)
D_{Friday}	0.1973***	(0.0043)	0.2050***	(0.0041)
$D_{Saturday}$	0.1540***	(0.0042)	0.1574***	(0.0041)
D_{Sunday}	0.1765***	(0.0045)	0.1881***	(0.0045)
D_{PH1}	0.0033***	(0.0006)	0.0042***	(0.0006)
D_{PH2}			0.0859***	(0.0026)
D_{PH3}			0.0917***	(0.0025)
D_{PB}	-0.5736***	(0.0188)	-0.6058***	(0.0192)
D_{TO}	-0.0750***	(0.0043)	-0.0365***	(0.0054)
Constant	0.9042***	(0.0039)	0.9312***	(0.0047)
Observations	617,076		1,005,634	
R^2	0.7005		0.6805	
R^2 within	0.1189		0.0947	

The estimation is performed by using a GMM. Dependent variable is daily frequency and is expressed in logarithm. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4
Frequencies groups 1 - fixed-effect panel regression results.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{Monday}	0.1487***	(0.0055)	0.1454***	(0.0055)
$D_{Wednesday}$	-0.0202***	(0.0028)	-0.0248***	(0.0027)
$D_{Thursday}$	0.1092***	(0.0049)	0.1039***	(0.0050)
D_{Friday}	0.2157***	(0.0061)	0.2202***	(0.0063)
$D_{Saturday}$	0.1828***	(0.0055)	0.1793***	(0.0056)
D_{Sunday}	0.2120***	(0.0060)	0.2186***	(0.0063)
D_{PH1}	0.0066***	(0.0008)	0.0083***	(0.0008)
D_{PH2}			0.0746***	(0.0030)
D_{PH3}			0.0736***	(0.0029)
D_{TO}	-0.0921***	(0.0040)	-0.0854***	(0.0055)
Constant	0.6778***	(0.0042)	0.7227***	(0.0056)
Observations	291,532		492,075	
R^2	0.8937		0.8478	
R^2 within	0.1368		0.1002	

The estimation is performed by using a GMM. Dependent variable is daily frequency and is expressed in logarithm. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

to the industry average and the effect of the takeover on the remaining operators' average supply of bus rides, respectively. All time-invariant heterogeneity between different routes is absorbed by the route fixed-effect α_i.

Table 3 shows the results of regression (9) for the takeover-period and the post-takeover-period. Again, there is a distinction between a short-term and a long-term perspective. In both cases, the pre-takeover-period is September 2016 – October 2016. From the short-term-perspective the period November 2016 – December 2016 is the post-takeover period and from the long-term-perspective that period is November 2016 – December 2016 and September 2017 – December 2017.

The coefficient of the takeover-dummy $D_{TO,t}$ shows that the daily average number of trips per route of the remaining operators significantly decreased. This is in line with the observations explained above. As a next step, we focus on data from group 1, i.e., the group where Flixbus is the sole operator before and after the takeover. This allows us to evaluate the service provision of Flixbus after the takeover.

Table 4 presents the results of our regressions. Note that, similar to Section 4.3, these figures only contain data on Flixbus. One can see that, on average, Flixbus decreased the daily number of trips per route by about 9.21% in the year of the takeover (short-term-perspective) and by 8.54% in the year after the takeover (long-term-perspective). This demonstrates that Flixbus significantly reduced the frequencies on the respective routes and that this effect was long-lasting.

Given that the takeover was announced in advance, Flixbus might have adjusted its strategy already before the observation period. However, even if frequencies were reduced before the takeover the above results show that there still is a significant decrease in frequencies due to the takeover. Thus, one would expect our results to underestimate the reduction in frequencies resulting from the takeover.

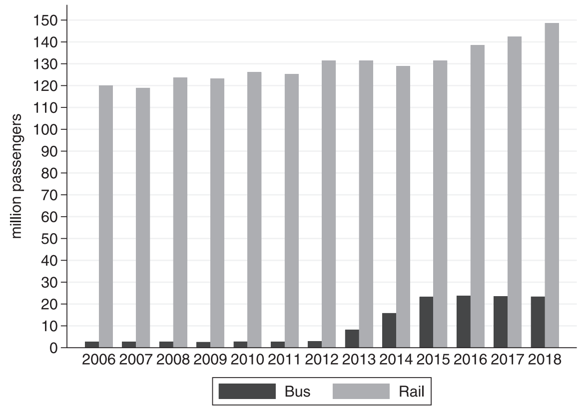


Fig. 5. Number of interurban bus and rail passengers from 2006 to 2018. Source: Federal Statistical Office of Germany. see https://bit.ly/3h7QpnS and https://bit.ly/3BabzvI (last accessed February 10, 2022).

We conduct the same robustness checks for the analyses in this section as for the analyses of prices (see above). One can see that our results are robust. For further detail, see Appendix A.

4.5. Intermodal competition

As a last step of our empirical analysis we examine the role of intermodal competition. Traveling by train can be considered a close substitute to interurban bus transport.[28] That there exists intermodal competition between train and interurban bus transport was already pointed out by, e.g., Beestermöller (2017).

Fig. 5 shows that after the initial increase in the demand for interurban bus travel from 2013 to 2015, the number of passengers requesting interurban bus services has remained almost constant from 2015 to 2018. One can also see that between 2013 and 2014 the number of passengers traveling by train decreased mildly, which indicates that there may exist a substitution pattern between interurban bus and train travel. However, from 2015 to 2018 there has been a relatively strong increase in the number of railway passengers again. These figures indicate that there was an overall increase in demand for transport in Germany, which started in 2013. In particular, these developments are not consistent with a decrease in demand for interurban bus travel that would explain the decrease in prices and frequencies after the takeover. This is analyzed in more detail in Appendix A.

To control for the effects of railway connections on the prices in the interurban bus industry, we perform a fixed-effects panel regression. The structural equation of this regression takes the following form:

$$\ln p_{i,t} = X_i'\theta_1 + \gamma_1 D_{RailDirect,i} + \gamma_2 D_{RailChange,i} + \gamma_3 D_{Group1,i} + \eta_t + \zeta_j + \mu_{i,j,t}. \quad (10)$$

The dependent variable in Eq. (10) is the logarithm of daily average prices per km for each route i provided at day t. Only data for Flixbus are included in this regression. The matrix X_i contains socio-demographic covariates for each route, i.e., population and the share of population under 25 years of age.[29] We also control for the duration of the journey (in minutes), distance (in meters) and squared distance for each route.[30] The dummy $D_{RailDirect,i}$ takes the value 1 for each route for which there is a direct railway connection available and is 0 otherwise. The dummy variable $D_{RailChange,i}$ takes the value 1 for routes for which there only is a railway connection available with one changeover. Otherwise, $D_{RailChange,i}$ is zero. A comparable approach was chosen by Dürr et al. (2016). The dummy variable $D_{Group1,i}$ is 1 for all routes in group 1 (see chapter 4.2) and 0 for all other routes. All time-variant heterogeneity is absorbed by the time fixed effect η_t. Heterogeneity between providers is absorbed by the provider fixed-effect ζ_j.

Table 5 presents the results of the regression of Eq. (10). The coefficients for $D_{RailDirect}$ and $D_{RailChange}$, are significantly negative. Hence, the availability of an alternative railway connection has a negative impact on prices in the interurban bus industry. This indicates that consumers perceive railway travel as a close substitute to interurban bus travel. Thus, the market size for interurban bus services is to a large extent limited by train connections. The consumers' outside option is therefore affected by railway passenger transport. Thus, we find strong empirical evidence for our corresponding assumptions in Section 3. Notably, one could argue that intermodal competition (Flixbus vs. train) is stronger than intramodel competition (Flixbus vs. other interurban bus suppliers). Our results indicate that the number of intramodal competitors (group 1

[28] On some routes such as Cologne - Berlin as well as Munich - Berlin, domestic flights are available. As already described above, we do not take into account this type of transport in our analyses because - if not booked several weeks in advance - fares are usually far above those for bus transport.

[29] The values are summed up for the origin and destination.

[30] Given the (time-invariant) variables of interest, i.e., $D_{RailDirect,i}$, $D_{RailChange,i}$ and $D_{Group1,i}$, we can not use route fixed-effects in this regression. Therefore, we include socio-demographic variables in order to control for (time-invariant) heterogeneity.

Table 5
Price influences of railway - fixed-effect panel regression results.

Model	Price per km	
Population	-1.38e-08***	(2.55e-09)
Population < 25	-1.0735***	(0.1797)
Distance	-1.87e-06***	(8.31e-08)
$Distance^2$	1.34e-12***	(8.22e-14)
Duration	3.05e-05	(0.0000)
$D_{RailDirect}$	-0.0590***	(0.0077)
$D_{RailChange}$	-0.0440***	(0.0071)
D_{GROUP1}	0.0172***	(0.0048)
Constant	2.4414***	(0.0493)
Observations	277,762	
R^2	0.4541	
R^2 within	0.3876	

Dependent variable is daily average price per kilometer expressed in logarithm. The regression includes time fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

vs. others) has a much smaller impact on prices than whether there is also a parallel railway connection available on the respective route.[31]

5. Conclusion

Our analyses indicate that there was entry deterrence by means of product proliferation in the German interurban bus industry. Based on a theoretical model, we expect that such a strategy of product proliferation is consistent with a decrease in the supply of bus rides of the incumbent (Flixbus) per route and day after the threat of entry of a potential entrant (Postbus) was eliminated, i.e., after the incumbent took over the entrant. We indeed identify a significant decrease in the daily supply of bus rides per route and day offered by Flixbus.

We also identify an increase in the industry-wide price level on average. This is because Postbus, who left the market, was a low-price carrier. The exit of that operator unambiguously increased the industry-wide average price level and decreased the average supply of bus rides per route and day. Controlling for this effect, the case at hand constitutes a particularly interesting example of product proliferation because the incumbent firm did not noticeably increase prices after the threat of entry was eliminated.

Our results have implications for merger analysis. We have presented an example where non-increasing price levels after a takeover can be interpreted as a sign of market power. Thus, despite the merging firms charging constant or even decreasing prices post-merger, the merger might have caused a decline in consumer surplus because of the decrease in variety. Decreasing prices alone are thus not necessarily an indication that a merger increases welfare. Welfare effects could be quantified if data on the exact number of passengers, i.e., on market demand, were available. Although we cannot quantify the welfare effects of the merger, our results emphasize the importance of variety, especially in differentiated goods market such as passenger transport.

Appendix A

As a robustness check for potential endogeneity, we apply the heteroscedasticity based IV approach as suggested by Lewbel (2012). This method can be used to identify structural parameters in models with endogenous regressors if no traditional identifying information, e.g., external instruments, are available. In this context, identification is achieved using regressors that are uncorrelated with the product of heteroscedastic errors. Instruments are constructed as simple functions of the data of the model.[32] We use this method to instrument for the potentially endogenous variable $D_{TO,t}$. The following tables contain the estimation results of the corresponding Lewbel IV regressions (based on Eq. (8)) for prices and frequencies in group 1, respectively. The results for both the takeover-period and the post-takeover-period are presented below.

The Kleibergen-Paap statistic suggests that the instruments are sufficiently strong because the critical values for weak identification are exceeded by far.[33] These results do not indicate the presence of endogeneity.

Fig. 5 shows no clear trend which could serve as an alternative explanation for our findings. To check for robustness of our analyses against the hypothesis of an exogenous trend, we estimate a difference-in-differences regression. In doing

[31] The sign of the coefficients of the socio-demographic covariates and of the duration as well as distance are in line with the ones reported by Dürr et al. (2016). We include these variables in our regression only to control for heterogeneity. The reader is thus directed to their article for a detailed discussion.

[32] For a short introduction to this method see, e.g., Baum et al. (2012)

[33] The null hypothesis of the Kleibergen-Paap test is that the structural equation is under-identified (i.e., the rank condition fails). Critical values are taken from Stock and Yogo (2002). As a rule of thumb, a value for the test statistic above ten indicates identification of the model.

Table A.1
Lewbel's IV regression - prices group 1.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{TO}	-0.0155***	(0.0005)	-0.0027***	(0.0005)
D_{Monday}	0.0510***	(0.0010)	0.0181***	(0.0008)
$D_{Wednesday}$	-0.0123***	(0.0008)	-0.0166***	(0.0008)
$D_{Thursday}$	0.0169***	(0.0008)	-0.0028***	(0.0008)
D_{Friday}	0.1371***	(0.0010)	0.1236***	(0.0009)
$D_{Saturday}$	0.0473***	(0.0009)	0.0412***	(0.0008)
D_{Sunday}	0.2015***	(0.0011)	0.1910***	(0.0009)
D_{PH1}	0.0523***	(0.0010)	0.0510***	(0.0011)
D_{PH2}			0.0986***	(0.0011)
D_{PH3}			0.1263***	(0.0012)
Observations	291,532		492,075	
R^2	0.2115		0.2108	
Weak identification test:				
Kleibergen-Paap rk Wald F statistic		3.54e+05		2.62e+06
Stock-Yogo (2002) critical values:		5% max IV rel bias		19.86
		10% max IV size		31.50

D_{TO} is instrumented using Lewbel (2012) heteroscedasticity based IV method. The estimation is performed by using a 2-step GMM. Dependent variable is daily average price per kilometer and is expressed in logarithms. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.2
Lewbel's IV regression - frequencies group 1.

Model	short-term		long-term	
pre-takeover	Sep. - Oct. 2016		Sep. - Oct. 2016	
post-takeover	Nov. - Dec. 2016		Nov. - Dec. 2016 & Sep. - Dec. 2017	
D_{TO}	-0.0880***	(0.0010)	-0.0837***	(0.0010)
D_{Monday}	0.1470***	(0.0018)	0.1444***	(0.0017)
$D_{Wednesday}$	-0.0151***	(0.0021)	-0.0241***	(0.0020)
$D_{Thursday}$	0.1073***	(0.0018)	0.1031***	(0.0017)
D_{Friday}	0.2096***	(0.0018)	0.2195***	(0.0017)
$D_{Saturday}$	0.1757***	(0.0018)	0.1784***	(0.0017)
D_{Sunday}	0.2051***	(0.0018)	0.2181***	(0.0017)
D_{PH1}	0.0065***	(0.0011)	0.0075***	(0.0017)
D_{PH2}			0.0715***	(0.0015)
D_{PH3}			0.0720***	(0.0015)
Observations	291,532		492,075	
R^2	0.1365		0.1002	
Weak identification test:				
Kleibergen-Paap rk Wald F statistic		3.54e+05		2.62e+06
Stock-Yogo (2002) critical values:		5% max IV rel bias		19.86
		10% max IV size		31.50

D_{TO} is instrumented using Lewbel (2012) heteroscedasticity based IV method. The estimation is performed by using a 2-step GMM. Dependent variable is daily frequency and is expressed in logarithms. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * p < 0.05, ** p < 0.01, *** p < 0.001.

so, we run regressions (8) and (9) using all routes in group 1 as the treatment group and all routes in group 4 as the control group. The estimation is run using data on Flixbus' prices and frequencies. In order to account for a potential bias that stems from the exit of another operator, we keep only such routes in the control group that where served by other operators throughout the entire observation period. Groups 3 and 6 as well as 2 and 5 are excluded from the analysis because they could have been affected by the merger.

The estimation results are shown in Tables A.3 and A.4, respectively. The results do not indicate that our results are driven by an exogenous trend. In order to provide evidence that the difference-in-differences-specification provides robust results, we have applied two robustness checks. The robustness checks should provide evidence that the parallel trends assumption holds. First, Figs. A.1 and A.2 show a graphical comparison of the prices and frequencies in the treatment and control group pre-merger. Second, we run an event study where we analyze the weekly difference-in-differences in order to check if there is no significantly different evolution between the two groups in the pre-treatment period. We interpret the results in a way that parallel trends can be assumed, even-though the parallel-trends assumption seems to be more robust for frequencies than for prices. Since the specifications in the main text also do not show a clear indication of a strong change in prices, the results of the event study do not seem to be surprising. It should be mentioned that the difference-

Table A.3
Differences-in-differencs regression - prices.

Model	Model 1		Model 2		Model 3	
D_{Monday}	0.0133***	(0.0017)	0.0214***	(0.0011)		
$D_{Wednesday}$	-0.0149***	(0.0013)	-0.0165***	(0.0007)		
$D_{Thursday}$	-0.0067***	(0.0017)	-0.0017	(0.0011)		
D_{Friday}	0.1161***	(0.0024)	0.1262***	(0.0019)		
$D_{Saturday}$	0.0321***	(0.0020)	0.0432***	(0.0013)		
D_{Sunday}	0.1809***	(0.0030)	0.1927***	(0.0027)		
D_{PH1}	0.0527***	(0.0010)	0.0533***	(0.0010)		
D_{PH2}	0.0961***	(0.0015)	0.0990***	(0.0014)		
D_{PH3}	0.1221***	(0.0019)	0.1250***	(0.0018)		
D_{2017}	0.0066***	(0.0021)	0.0046*	(0.0024)		
Population	-7.72e-08***	(4.49e-09)	2.88e-07*	(1.63e-07)	2.77e-07*	(1.64e-07)
Population < 25	-1.8267***	(0.3495)	-1.5120*	(0.8518)	-1.3661	(0.8517)
D_{TO}	0.0303***	(0.0079)	0.0282***	(0.0079)		
D_{GROUP1}	0.0096	(0.0173)				
$D_{GROUP1} * D_{TO}$	-0.0324***	(0.0083)	-0.0335***	(0.0083)	-0.0344***	(0.0083)
Constant	2.2104***	(0.0899)	1.8613***	(0.2380)	1.9337***	(0.2397)
Observations	515,562		515,562		515,562	
R^2	0.1222		0.7350		0.7717	
R^2 within	0.1222		0.2171		0.0007	
Time Fixed	No		No		Yes	
Route Fixed	No		Yes		Yes	

The estimation is performed by using a GMM. Dependent variable is daily average price per kilometer and is expressed in logarithms. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * p < 0.05, ** p < 0.01, *** p < 0.001.

Table A.4
Differences-in-differencs regression - frequencies.

Model	Model 1		Model 2		Model 3	
D_{Monday}	0.1109***	(0.0059)	0.1459***	(0.0053)		
$D_{Wednesday}$	-0.0283***	(0.0038)	-0.0252***	(0.0026)		
$D_{Thursday}$	0.0709***	(0.0059)	0.1056***	(0.0048)		
D_{Friday}	0.1767***	(0.0071)	0.2230***	(0.0060)		
$D_{Saturday}$	0.1349***	(0.0065)	0.1809***	(0.0054)		
D_{Sunday}	0.1696***	(0.0071)	0.2208***	(0.0060)		
D_{PH1}	0.0089***	(0.0009)	0.0086***	(0.0008)		
D_{PH2}	0.0610***	(0.0026)	0.0678***	(0.0023)		
D_{PH3}	0.0606***	(0.0025)	0.0681***	(0.0022)		
D_{2017}	0.0172***	(0.0055)	0.0281***	(0.0067)		
Population	7.85e-08***	(1.65e-08)	7.29e-07	(6.31e-07)	7.47e-07	(6.32e-07)
Population < 25	-1.2786	(0.7841)	-3.7870	(2.7890)	-3.6399	(2.7918)
D_{TO}	0.0490*	(0.0256)	0.0249	(0.0253)		
D_{GROUP1}	-0.6125***	(0.0856)				
$D_{GROUP1} * D_{TO}$	-0.1427***	(0.0266)	-0.1276***	(0.0263)	-0.1322***	(0.0264)
Constant	1.6261***	(0.2118)	1.1408	(0.7920)	1.2654	(0.7965)
Observations	515,562		515,562		515,562	
R^2	0.0651		0.8662		0.8706	
R^2 within	0.0651		0.1072		0.0026	
Time Fixed	No		No		Yes	
Route Fixed	No		Yes		Yes	

The estimation is performed by using a GMM. Dependent variable is daily frequency and is expressed in logarithms. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * p < 0.05, ** p < 0.01, *** p < 0.001.

in-differences-analysis and the results of the parallel trends tests might be biased by the market exit of another operator (BerlinLinienBus) and need to be treated carefully. This is why this specification is presented in the Appendix rather than in the main text.

We also check whether our results are driven by intra-year fluctuations (i.e., frequencies in September and October could always be higher than in other months due to some (unobservable) seasonal trends). To identify potential seasonal patterns, we estimate regressions (8) and (9) using only data for September and October (months that are represented in the pre- and post-merger period). Corresponding results are stated in Tables A.7 and A.8. The results of these estimations do not indicate that our findings in the main text can be explained by seasonal patterns.

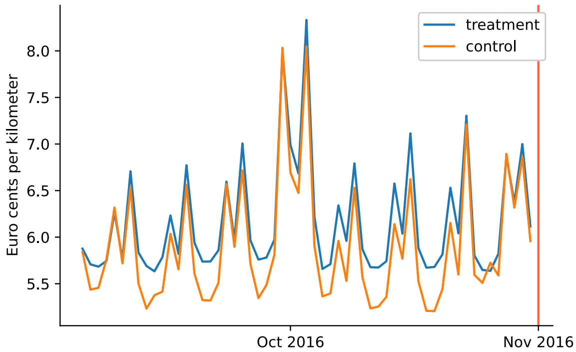


Fig. A.1. Differences-in-differencs - average weekly prices g1 vs g4

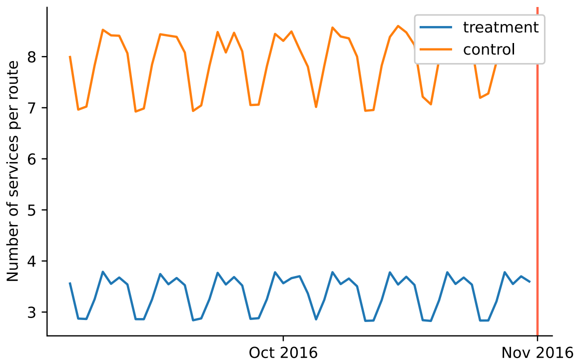


Fig. A.2. Differences-in-differencs - average weekly frequencies g1 vs g4

Table A.5
Event study - prices.

Model	Price per km	
Population	2.60e-07	(1.64e-07)
Population < 25	-1.2963	(0.8466)
$D_{Lead\ 8}$	-0.0114	(0.0072)
$D_{Lead\ 7}$	0.0169**	(0.0070)
$D_{Lead\ 6}$	0.0205***	(0.0069)
$D_{Lead\ 5}$	0.0132*	(0.0080)
$D_{Lead\ 4}$	0.0248***	(0.0080)
$D_{Lead\ 3}$	0.0326***	(0.0076)
$D_{Lead\ 2}$	0.0306***	(0.0067)
$D_{Lag\ 0}$	0.0072	(0.0081)
$D_{Lag\ 1}$	0.0187***	(0.0068)
$D_{Lag\ 2}$	0.0187**	(0.0075)
$D_{Lag\ 3}$	0.0232***	(0.0077)
$D_{Lag\ 4}$	0.0157**	(0.0077)
$D_{Lag\ 5}$	0.0066	(0.0078)
$D_{Lag\ 6}$	-0.0389***	(0.0139)
$D_{Lag\ 7}$	-0.0309**	(0.0131)
$D_{Lag\ 8}$	-0.0463***	(0.0128)
$D_{Lag\ 9}$	-0.0561***	(0.0129)
$D_{Lag\ 10}$	-0.0394***	(0.0126)
$D_{Lag\ 11}$	-0.0185	(0.0127)
$D_{Lag\ 12}$	-0.0188	(0.0132)
$D_{Lag\ 13}$	-0.0214*	(0.0125)
$D_{Lag\ 14}$	-0.0278*	(0.0154)
$D_{Lag\ 15}$	-0.0321***	(0.0125)
$D_{Lag\ 16}$	-0.0284**	(0.0128)
$D_{Lag\ 17}$	-0.0257**	(0.0128)
$D_{Lag\ 18}$	-0.0321**	(0.0127)
$D_{Lag\ 19}$	-0.0432***	(0.0133)
Constant	1.9145***	(0.2385)
Observations	515,562	
R^2	0.7719	
R^2 within	0.0017	
Time Fixed	Yes	
Route Fixed	Yes	

Dependent variable is daily average price per kilometer expressed in logarithm. The regression includes time and route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.6
Event study - frequencies.

Model	Daily frequency	
Population	6.94e-07	(6.27e-07)
Population < 25	-3.4230	(2.7861)
$D_{Lead\ 8}$	0.0219***	(0.0081)
$D_{Lead\ 7}$	0.0081	(0.0072)
$D_{Lead\ 6}$	0.0124*	(0.0075)
$D_{Lead\ 5}$	0.0061	(0.0044)
$D_{Lead\ 4}$	0.0005	(0.0046)
$D_{Lead\ 3}$	0.0050	(0.0040)
$D_{Lead\ 2}$	-0.0057**	(0.0024)
$D_{Lag\ 0}$	-0.0162***	(0.0054)
$D_{Lag\ 1}$	-0.0202	(0.0195)
$D_{Lag\ 2}$	-0.0219	(0.0218)
$D_{Lag\ 3}$	-0.0135	(0.0223)
$D_{Lag\ 4}$	-0.0286	(0.0256)
$D_{Lag\ 5}$	-0.0080	(0.0250)
$D_{Lag\ 6}$	-0.1931***	(0.0329)
$D_{Lag\ 7}$	-0.1950***	(0.0327)
$D_{Lag\ 8}$	-0.1823***	(0.0325)
$D_{Lag\ 9}$	-0.1736***	(0.0326)
$D_{Lag\ 10}$	-0.1800***	(0.0334)
$D_{Lag\ 11}$	-0.1928***	(0.0330)
$D_{Lag\ 12}$	-0.1813***	(0.0326)
$D_{Lag\ 13}$	-0.1871***	(0.0325)
$D_{Lag\ 14}$	-0.1834***	(0.0330)
$D_{Lag\ 15}$	-0.1538***	(0.0341)
$D_{Lag\ 16}$	-0.1331***	(0.0342)
$D_{Lag\ 17}$	-0.1494***	(0.0343)
$D_{Lag\ 18}$	-0.1437***	(0.0348)
$D_{Lag\ 19}$	-0.1456***	(0.0344)
Constant	1.2472	(0.7930)
Observations	515,562	
R^2	0.8709	
R^2 within	0.0048	
Time Fixed	Yes	
Route Fixed	Yes	

Dependent variable is daily frequency expressed in logarithm. The regression includes time and route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.7
September and October only regression - prices.

Model	Price per km	
D_{Monday}	0.0271***	(0.0012)
$D_{Wednesday}$	-0.0309***	(0.0009)
$D_{Thursday}$	-0.0054***	(0.0012)
D_{Friday}	0.1408***	(0.0021)
$D_{Saturday}$	0.0633***	(0.0016)
D_{Sunday}	0.1757***	(0.0027)
D_{PH1}	0.0537***	(0.0010)
D_{PH2}	0.0792***	(0.0012)
D_{PH3}	0.1128***	(0.0021)
D_{TO}	0.0182***	(0.0021)
Constant	1.7040***	(0.0016)
Observations	295,188	
R^2	0.7237	
R^2 within	0.2030	

The estimation is performed by using a GMM. Dependent variable is daily average price per kilometer and is expressed in logarithms. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.8
September and October only regression - frequencies.

Model	Daily frequency	
D_{Monday}	0.1215***	(0.0053)
$D_{Wednesday}$	-0.0287***	(0.0026)
$D_{Thursday}$	0.0797***	(0.0046)
D_{Friday}	0.1790***	(0.0060)
$D_{Saturday}$	0.1395***	(0.0054)
D_{Sunday}	0.1751***	(0.0060)
D_{PH1}	0.0082***	(0.0008)
D_{PH2}	0.0238***	(0.0013)
D_{PH3}	0.0166***	(0.0013)
D_{TO}	-0.0327***	(0.0067)
Constant	0.7480***	(0.0051)
Observations	295,188	
R^2	0.8796	
R^2 within	0.0766	

The estimation is performed by using a GMM. Dependent variable is daily frequency and is expressed in logarithms. The regression includes route fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.9
Influence of railway on frequency - fixed-effect panel regression results.

Model	Daily frequency	
Population	2.51e-07***	(1.38e-08)
Population < 25	-2.3893***	(0.6500)
Distance	-3.36e-06***	(3.70e-07)
$Distance^2$	1.08e-12***	(3.29e-13)
Duration	0.0015***	(0.0003)
$D_{RailDirect}$	-0.1021***	(0.0343)
$D_{RailChange}$	-0.2789***	(0.0294)
D_{GROUP1}	-0.5211***	(0.0244)
Constant	2.3399***	(0.1697)
Observations	277,762	
R^2	0.2206	
R^2 within	0.2180	

Dependent variable is daily frequency expressed in logarithm. The regression includes time fixed-effects. Cluster-robust standard errors (clustered on route level) are presented in parentheses. Statistics are significant for * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.10
Pre-takeover period 2016 - summary statistics

Group	Variable	count	mean	std	min	25%	50%	75%	max
1	*Price*	193989	6.08647	2.12712	2.03364	4.73219	5.48853	6.84655	26.0706
	Frequency	193989	3.11339	3.95628	1	1	2	4	64
	Distance	193989	301661	177938	37700	145803	272773	433037	934286
	Distance^2	193989	1.22661e+11	1.3051e+11	1.42129e+09	2.12585e+10	7.44051e+10	1.87521e+11	8.7289e+11
	Duration	193989	305.906	181.418	30	150	272.5	440	980
	D_{PB}	193989	0	0	0	0	0	0	0
	D_{TO}	193989	0	0	0	0	0	0	0
	D_{PH}	193989	0.247163	0.431364	0	0	0	0	1
	$D_{RailDirect}$	193989	0.199357	0.399518	0	0	0	0	1
	$D_{RailChange}$	193989	0.530417	0.499075	0	0	1	1	1
	Population	193989	702727	927944	6983	147958	377040	699418	3.85245e+06
	Population <25	193989	0.24654	0.0176971	0.175318	0.235575	0.244718	0.255132	0.315878
	PPP	193989	1.51095e+10	1.8941e+10	1.41709e+08	3.19895e+09	8.23012e+09	1.60207e+10	7.6615e+10
2	*Price*	12838	4.45694	2.11665	1.06995	3.25019	4.05934	5.06952	20.917
	Frequency	12838	1.37942	0.95513	1	1	1	1	7
	Distance	12838	344318	181845	20059	203531	312059	460597	751082
	Distance^2	12838	1.5162e+11	1.41682e+11	4.02363e+08	4.14249e+10	9.73808e+10	2.1215e+11	5.64124e+11
	Duration	12838	344.599	189.038	20	190	315	465	837.5
	D_{PB}	12838	1	0	1	1	1	1	1
	D_{TO}	12838	0	0	0	0	0	0	0
	D_{PH}	12838	0.24334	0.429115	0	0	0	0	1
	$D_{RailDirect}$	12838	0.335488	0.472179	0	0	0	1	1
	$D_{RailChange}$	12838	0.448434	0.497353	0	0	0	1	1
	Population	12838	516838	337692	14637	257993	534535	657442	1.58779e+06
	Population <25	12838	0.24425	0.018535	0.193105	0.233894	0.244085	0.254085	0.296087
	PPP	12838	1.16381e+10	9.0436e+09	3.44534e+08	5.22367e+09	1.08832e+10	1.42662e+10	4.46217e+10
3	*Price*	47463	4.95502	1.67994	1.06188	3.81491	4.76933	5.85451	18.4858
	Frequency	47463	2.69422	2.50671	1	1	2	3	20
	Distance	47463	330396	161768	67170	198527	315371	442140	724585
	Distance^2	47463	1.3533e+11	1.22026e+11	4.51181e+09	3.9413e+10	9.94589e+10	1.95488e+11	5.25023e+11
	Duration	47463	323.828	158.589	55	195	308.75	430	827.5
	D_{PB}	47463	0.49211	0.499943	0	0	0	1	1
	D_{TO}	47463	0	0	0	0	0	0	0
	D_{PH}	47463	0.245855	0.430597	0	0	0	0	1
	$D_{RailDirect}$	47463	0.486505	0.499823	0	0	0	1	1
	$D_{RailChange}$	47463	0.481217	0.499652	0	0	0	1	1
	Population	47463	1.01656e+06	854967	79103	530926	743250	1.23337e+06	3.86441e+06
	Population <25	47463	0.246809	0.0140934	0.190965	0.238272	0.246045	0.253238	0.312145
	PPP	47463	2.23526e+10	1.78147e+10	1.5761e+09	1.12505e+10	1.64537e+10	2.80177e+10	7.57188e+10
4	*Price*	62511	6.33407	4.44542	1.42897	4.55645	5.26809	6.70023	52.8983
	Frequency	62511	5.13287	7.3611	1	2	3	5	66
	Distance	62511	269442	149379	50276	150839	239327	363940	809540
	Distance^2	62511	9.49126e+10	1.03981e+11	2.52768e+09	2.27524e+10	5.72774e+10	1.32452e+11	6.55355e+11
	Duration	62511	254.871	148.002	45	141.25	215	337.5	990
	D_{PB}	62511	0	0	0	0	0	0	0
	D_{TO}	62511	0	0	0	0	0	0	0
	D_{PH}	62511	0.246005	0.430685	0	0	0	0	1
	$D_{RailDirect}$	62511	0.380909	0.485614	0	0	0	1	1
	$D_{RailChange}$	62511	0.494249	0.499971	0	0	0	1	1
	Population	62511	1.20878e+06	1.26536e+06	18436	286856	675635	1.49969e+06	4.122e+06
	Population <25	62511	0.245779	0.016618	0.17757	0.233857	0.24342	0.252376	0.317721
	PPP	62511	2.54456e+10	2.49391e+10	3.2311e+08	5.98766e+09	1.49348e+10	4.28182e+10	8.01632e+10
5	*Price*	11396	5.78371	4.05593	1.06987	3.60475	4.91033	6.46573	38.3586
	Frequency	11396	1.86618	1.36041	1	1	1	2	9
	Distance	11396	324470	179321	82302	159136	298409	482662	700652
	Distance^2	11396	1.37434e+11	1.32851e+11	6.77362e+09	2.53243e+10	8.90479e+10	2.32963e+11	4.90913e+11
	Duration	11396	331.843	205.207	55	170	285	480	905
	D_{PB}	11396	0.523342	0.499477	0	0	1	1	1
	D_{TO}	11396	0	0	0	0	0	0	0
	D_{PH}	11396	0.247367	0.431501	0	0	0	0	1
	$D_{RailDirect}$	11396	0.373991	0.483882	0	0	0	1	1
	$D_{RailChange}$	11396	0.424272	0.494254	0	0	0	1	1
	Population	11396	545590	407338	59796	237492	524670	669148	1.9507e+06
	Population <25	11396	0.240176	0.0162616	0.188578	0.232944	0.241632	0.250597	0.273982
	PPP	11396	1.23237e+10	1.05324e+10	1.42644e+09	4.91887e+09	1.00141e+10	1.61929e+10	4.61071e+10
6	*Price*	125141	5.06804	1.88017	1.0032	4.01542	4.79836	5.76104	37.7727
	Frequency	125141	4.80635	5.78085	1	1	3	6	56
	Distance	125141	317237	158897	53290	190678	297069	431319	777097
	Distance^2	125141	1.25887e+11	1.13243e+11	2.83982e+09	3.63581e+10	8.825e+10	1.86036e+11	6.0388e+11
	Duration	125141	299.278	157.34	50	170	275	415	845
	D_{PB}	125141	0.332113	0.470973	0	0	0	1	1
	D_{TO}	125141	0	0	0	0	0	0	0
	D_{PH}	125141	0.245555	0.430418	0	0	0	0	1
	$D_{RailDirect}$	125141	0.637241	0.480798	0	0	1	1	1
	$D_{RailChange}$	125141	0.310666	0.462768	0	0	0	1	1
	Population	125141	1.38555e+06	1.1175e+06	109606	652041	992952	1.70397e+06	5.39189e+06
	Population <25	125141	0.243149	0.0114444	0.18343	0.236556	0.24359	0.247379	0.293987
	PPP	125141	3.07735e+10	2.3102e+10	2.09204e+09	1.39878e+10	2.24216e+10	4.45735e+10	1.12437e+11

Table A.11
Post-takeover period 2016 - summary statistics

Group	Variable	count	mean	std	min	25%	50%	75%	max
1	*Price*	154765	5.91255	2.05195	2.32578	4.62558	5.34131	6.6127	61.1586
	Frequency	154765	3.15092	3.44519	1	1	2	4	64
	Distance	154765	306165	172733	37700	155367	283687	431985	934286
	$Distance^2$	154765	1.23574e+11	1.26649e+11	1.42129e+09	2.41389e+10	8.04783e+10	1.86611e+11	8.7289e+11
	Duration	154765	304.294	172.02	30	157.5	278.333	430	980
	D_{PB}	154765	0.00129874	0.0360148	0	0	0	0	1
	D_{TO}	154765	1	0	1	1	1	1	1
	D_{PH}	154765	0	0	0	0	0	0	0
	$D_{RailDirect}$	154765	0.308713	0.461964	0	0	0	1	1
	$D_{RailChange}$	154765	0.512086	0.499856	0	0	1	1	1
	Population	154765	831345	1.00057e+06	6983	193167	538658	837645	5.03913e+06
	Population < 25	154765	0.246049	0.0166712	0.175318	0.23519	0.244372	0.254007	0.317721
	PPP	154765	1.79707e+10	2.05633e+10	1.41709e+08	4.10306e+09	1.0786e+10	1.90631e+10	1.11783e+11
2	*Price*	49898	6.07023	2.76858	1.63067	4.61737	5.36691	6.68219	40.3777
	Frequency	49898	6.07543	7.88579	1	2	3	7	70
	Distance	49898	266833	148504	50276	147490	234488	366916	678268
	$Distance^2$	49898	9.3253e+10	9.75389e+10	2.52768e+09	2.17533e+10	5.49846e+10	1.34627e+11	4.60047e+11
	Duration	49898	263.067	164.688	45	135	217.5	360	2070
	D_{PB}	49898	0.00559141	0.0745671	0	0	0	0	1
	D_{TO}	49898	1	0	1	1	1	1	1
	D_{PH}	49898	0	0	0	0	0	0	0
	$D_{RailDirect}$	49898	0.607279	0.488361	0	0	1	1	1
	$D_{RailChange}$	49898	0.330294	0.470324	0	0	0	1	1
	Population	49898	1.1943e+06	1.02748e+06	24524	576033	903035	1.40939e+06	5.39189e+06
	Population < 25	49898	0.246915	0.0125219	0.207112	0.238359	0.245177	0.250717	0.301337
	PPP	49898	2.67803e+10	2.1237e+10	5.44051e+08	1.27503e+10	2.1163e+10	3.33182e+10	1.12437e+11
3	*Price*	1641	14.1623	10.345	2.9256	9.01984	11.6701	15.7977	57.8901
	Frequency	1641	4.78915	8.00301	1	1	3	4	45
	Distance	1641	312470	182778	61449	159485	279931	433589	683620
	$Distance^2$	1641	1.31025e+11	1.3342e+11	3.77598e+09	2.54355e+10	7.83614e+10	1.87999e+11	4.67336e+11
	Duration	1641	319.513	212.447	60	160	286.667	420	870
	D_{PB}	1641	0	0	0	0	0	0	0
	D_{TO}	1641	1	0	1	1	1	1	1
	D_{PH}	1641	0	0	0	0	0	0	0
	$D_{RailDirect}$	1641	0.221207	0.415186	0	0	0	0	1
	$D_{RailChange}$	1641	0.40585	0.491205	0	0	0	1	1
	Population	1641	1.08413e+06	1.15517e+06	98492	354234	590914	1.32089e+06	3.58771e+06
	Population < 25	1641	0.239294	0.0145817	0.188471	0.233314	0.239819	0.245102	0.274923
	PPP	1641	2.25897e+10	2.29015e+10	1.74868e+09	7.99579e+09	1.207e+10	2.93883e+10	6.99056e+10

Appendix B

Appendix B provides all proofs for Section 3.

Proof of Lemma. 1. Without loss of generality, consider the segment between $l_1 = 0$ and $l_1 = \frac{1}{n}$. Given the entrant's location is $l_2 \in \left(0, \frac{1}{n}\right)$, and given any prices p_1 and p_2, the consumer who is indifferent between acquiring the service at $l_1 = 0$ and l_2 is located at $\frac{l_2}{2} - \frac{p_1 - p_2}{2l_2}$. The ensuing demand for the entrant's outlet under consideration is thus $\frac{L}{n}\left(\frac{l_2}{2} + \frac{p_1 - p_2}{2l_2}\right)$, which yields that outlet's profit

$$\pi_{2,l_2} = \frac{L}{n}\left(\frac{l_2}{2} + \frac{p_1 - p_2}{2l_2}\right)p_2. \tag{1}$$

Using (1), the entrant has no incentive to move closer to $l_1 = 0$ if and only if $\frac{\partial \pi_{2,l_2}}{\partial l_2} \geq 0$, which is equivalent to

$$l_2^2 \geq p_1 - p_2. \tag{2}$$

Substituting the entrant's optimal price $p_2(p_1, l_2) = \frac{1}{2}\left(p_1 + l_2^2\right)$ which is obtained from the first order condition with respect to the price p_2, $\frac{\partial \pi_{2,l_2}}{\partial l_2}$ can be evaluated at $l_2 = \frac{1}{2n}$ (entry in the middle between $l_1 = 0$ and $l_1 = \frac{1}{n}$) by using (2):

$$\left(\frac{1}{2n}\right)^2 \geq p_1 - p_2(p_1, l_2) \Leftrightarrow p_1 \leq \frac{3}{4n^2}. \tag{3}$$

□

Proof of Lemma. 2. Substituting $p_M = u - \left(\frac{F^2}{16L^2}\right)^{\frac{1}{3}}$ (see (4)) into (5) (both in the main text) with $\epsilon \to 0$ yields

$$n_D \gtrapprox \frac{Lu - S}{F} - \left(\frac{L}{16F}\right)^{\frac{1}{3}} \equiv \underline{n}. \tag{4}$$

□

References

Aarhaug, J., Fearnley, N., 2016. Deregulation of the Norwegian long distance express coach market. Transp. Policy 46, 1–6.
Augustin, K., Gerike, R., Sanchez, M.J.M., Ayala, C., 2014. Analysis of intercity bus markets on long distances in an established and a young market: the example of the US and Germany. Res. Transp. Econ. 48, 245–254.
Bain, J.S., 1949. A note on pricing in monopoly and oligopoly. Am. Econ. Rev. 448–464.
Barroso, A., Giarratana, M.S., 2013. Product proliferation strategies and firm performance: The moderating role of product space complexity. Strateg. Manag. J. 34 (12), 1435–1452.
Bataille, M., Steinmetz, A., 2013. Intermodal competition on some routes in transportation networks: the case of inter urban buses and railways. DICE Discus. Paper 84.
Baum, C.F., Lewbel, A., Schaffer, M.E., Talavera, O., et al., 2012. Instrumental variables estimation using heteroskedasticity-based instruments. United Kingdom Stata Users Group Meetings, Vol. 7.
Bayus, B.L., Putsis, W.P., 1999. Product proliferation: An empirical analysis of product line determinants and market outcomes. Mark. Sci. 18 (2), 137–153.
Beestermöller, M., 2017. Striking Evidence? Demand Persistence for Inter-City Buses from German Railway Strikes. Münchener wirtschaftswissenschaftliche Beiträge 2.
Bergman, M.A., Rudholm, N., 2003. The relative importance of actual and potential competition: empirical evidence from the pharmaceuticals market. J. Ind. Econ. 51 (4), 455–467.
Beria, P., Grimaldi, R., Debernardi, A., Ferrara, E., Laurino, A., 2014. Spatial and scenario analyses of long distance coach transport in Italy..
Böckem, S., 1994. A generalized model of horizontal product differentiation. J. Ind. Econ. 287–298.
Böckers, V., Haucap, J., Heimeshoff, U., Thorwarth, S., 2015. Auswirkungen der Fernbusliberalisierung auf den Schienenpersonenfernverkehr. List Forum für Wirtschafts-und Finanzpolitik 41 (1), 75–90.
Bokhari, F., Yan, W., 2019. Product proliferation as entry deterrence: the case of UK pharmaceuticals.
Bonanno, G., 1987. Location choice, product proliferation and entry deterrence. Rev. Econ. Stud. 54 (1), 37–45.
Bronnenberg, B.J., Dhar, S.K., Dubé, J.P.H., 2009. Brand history, geography, and the persistence of brand shares. J. Polit. Econ. 117 (1), 87–115.
Bunch, D.S., Smiley, R., 1992. Who deters entry? evidence on the use of strategic entry deterrents. Rev. Econ. Stat. 509–521.
Bundesamt für Güterverkehr, 2015. Marktbeobachtung Güterverkehr. Marktanalyse des Fernbuslinienverkehrs 2014. Technical Report. Cologne, Germany
Bundesamt für Güterverkehr, 2016. Marktbeobachtung Güterverkehr. Marktanalyse des Fernbuslinienverkehrs 2015. Technical Report. Cologne, Germany
Cabral, L., Fishman, A., 2012. Business as usual: a consumer search theory of sticky prices and asymmetric price adjustment. Int. J. Ind. Organ. 30 (4), 371–376. doi:10.1016/j.ijindorg.2012.01.003.
Cookson, J.A., 2017. Anticipated entry and entry deterrence: evidence from the american casino industry. Manag. Sci. 64 (5), 2325–2344.
d'Aspremont, C., Gabszewicz, J.J., Thisse, J.F., 1979. On hotelling's" stability in competition". Econ. J. Econ. Soc. 1145–1150.
Dafny, L.S., 2005. Games hospitals play: entry deterrence in hospital procedure markets. J. Econ. Manag. Strateg. 14 (3), 513–542.
Dixit, A., 1980. The role of investment in entry-deterrence. Econ. J. 90 (357), 95–106.
Dixit, A., 1982. Recent developments in oligopoly theory. Am. Econ. Rev. 72 (2), 12–17.
Dürr, N., Hüschelrath, K., 2015. Competition in the German interurban bus industry: a snapshot two years after liberalization. Compet. Regul. Netw. Ind. 16 (3).
Dürr, N., Hüschelrath, K., Heim, S., 2016. Deregulation, competition, and consolidation: the case of the German interurban bus industry. J. Transp. Econ. Policy 50 (2).
Dürr, N.S., Hüschelrath, K., 2017. Patterns of entry and exit in the deregulated german interurban bus industry. Transp. Policy 59, 196–208.
Eaton, B.C., Lipsey, R.G., 1979. The theory of market pre-emption: the persistence of excess capacity and monopoly in growing spatial markets. Economica 46 (182), 149–158.

Economides, N., et al., 1986. Minimal and maximal product differentiation in hotellings duopoly. Econ. Lett. 21 (1), 67–71.
Ellison, G., Ellison, S.F., 2011. Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration. Am. Econ. J. Microecon. 3 (1), 1–36.
Farrell, J., Shapiro, C., 1990. Horizontal mergers: an equilibrium analysis. Am. Econ. Rev. 107–126.
Geroski, P.A., 1995. What do we know about entry? Int. J. Ind. Organ. 13 (4), 421–440.
Gilbert, R.J., 1989. Mobility barriers and the value of incumbency. Handb. Ind. Organ. 1, 475–535.
Goolsbee, A., Syverson, C., 2008. How do incumbents respond to the threat of entry? evidence from the major airlines. Q. J. Econ. 123 (4), 1611–1633.
Hagiu, A., 2006. Pricing and commitment by two-sided platforms. RAND J. Econ. 37 (3), 720–737.
Hall, E.A., 1990. An analysis of preemptive behavior in the titanium dioxide industry. Int. J. Ind. Organ. 8 (3), 469–484.
Judd, K.L., 1985. Credible spatial preemption. RAND J. Econ. 153–166.
King, A.A., Tucci, C.L., 2002. Incumbent entry into new market niches: the role of experience and managerial choice in the creation of dynamic capabilities. Manag. Sci. 48 (2), 171–186.
Knorr, A., Lueg-Arndt, A., 2016. Intercity bus deregulation in Germany–intramodal and intermodal effects after two years. Res. Transpo. Econ..
Kreps, D.M., Wilson, R., 1982. Reputation and imperfect information. J. Econ. Theory 27 (2), 253–279.
Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. J. Bus. Econ. Stat. 30 (1), 67–80.
Milgrom, P., Roberts, J., 1982. Limit pricing and entry under incomplete information: an equilibrium analysis. Econ. J. Econ. Soc. 443–459.
Moreno, A., Terwiesch, C., 2016. The effects of product line breadth: evidence from the automotive industry. Mark. Sci. 36 (2), 254–271.
Morton, F.M.S., 2000. Barriers to entry, brand advertising, and generic entry in the us pharmaceutical industry. int. J. ind. Organ. 18 (7), 1085–1104.
Neven, D.J., 1987. Endogenous sequential entry in a spatial model. Int. J. Ind. Organ. 5 (4), 419–434.
Regan, T.L., 2008. Generic entry, price competition, and market segmentation in the prescription drug market. Int. J. Ind. Organ. 26 (4), 930–948.
Salop, S.C., 1979. Monopolistic competition with outside goods. Bell J. Econ. 141–156.
Salop, S.C., 1979. Strategic entry deterrence. Am. Econ. Rev. 69 (2), 335–338.
Schmalensee, R., 1978. Entry deterrence in the ready-to-eat breakfast cereal industry. Bell J. Econ. 305–327.
Schmalensee, R., 1982. Product differentiation advantages of pioneering brands. Am. Econ. Rev. 72 (3), 349–365.
Shy, O., Stenbacka, R., 2006. Service hours with asymmetric distributions of ideal service time. Int. J. Ind. Organ. 24 (4), 764–771.
Spence, A.M., 1977. Entry, capacity, investment and oligopolistic pricing. Bell J. Econ. 534–544.
Stock, J.H., Yogo, M., 2002. Testing for weak instruments in linear IV regression. Technical Report.
Tirole, J., 1988. The Theory of Industrial Organization. MIT Press.
van de Velde, D., 2009. Long-distance Bus Services in Europe: Concessions or Free Market?, Vol. 21. OECD Publishing.
Weiman, D.F., Levin, R.C., 1994. Preying for monopoly? the case of southern bell telephone company, 1894-1912. J. Polit. Econ. 102 (1), 103–126.
White, P., Robbins, D., 2012. Long-term development of express coach services in Britain. Res. Transp. Econ. 36 (1), 30–38.
Wilson, R., 1992. Strategic models of entry deterrence. Handbook of Game Theory with Economic Applications 1, 305–329.

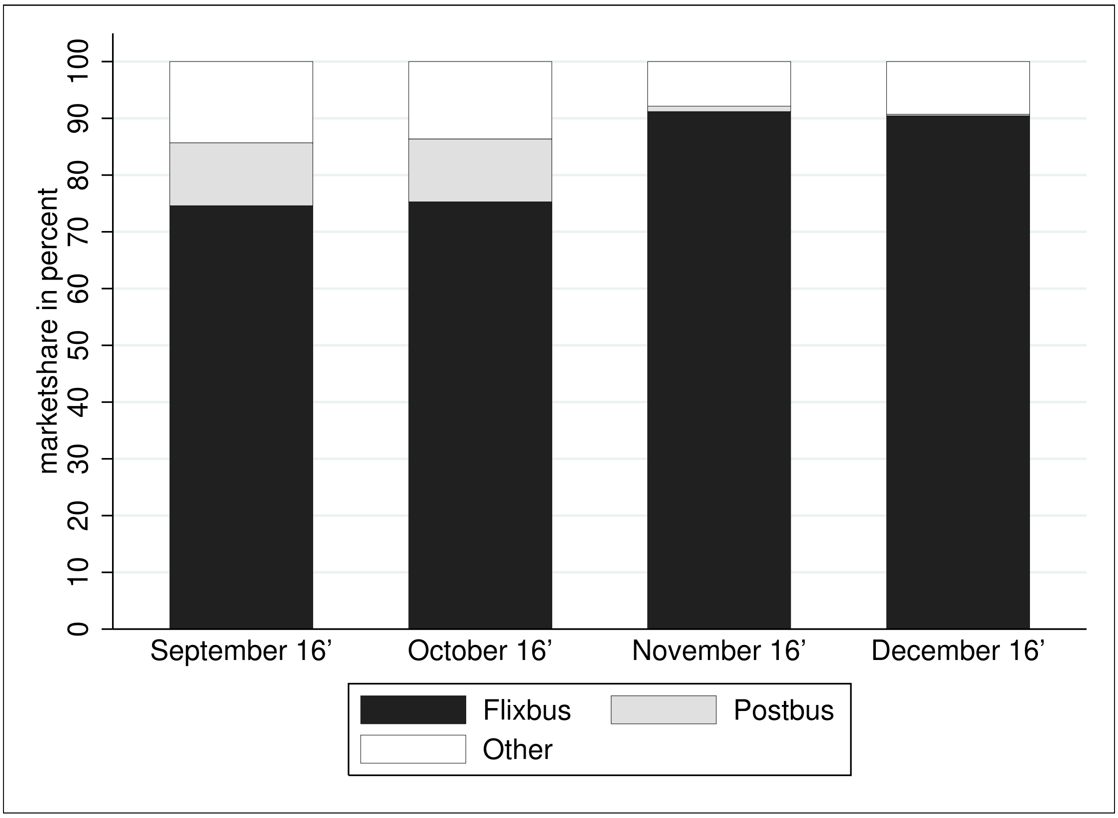


Figure 2.1.: Market shares before and after the takeover on November 1, 2016.

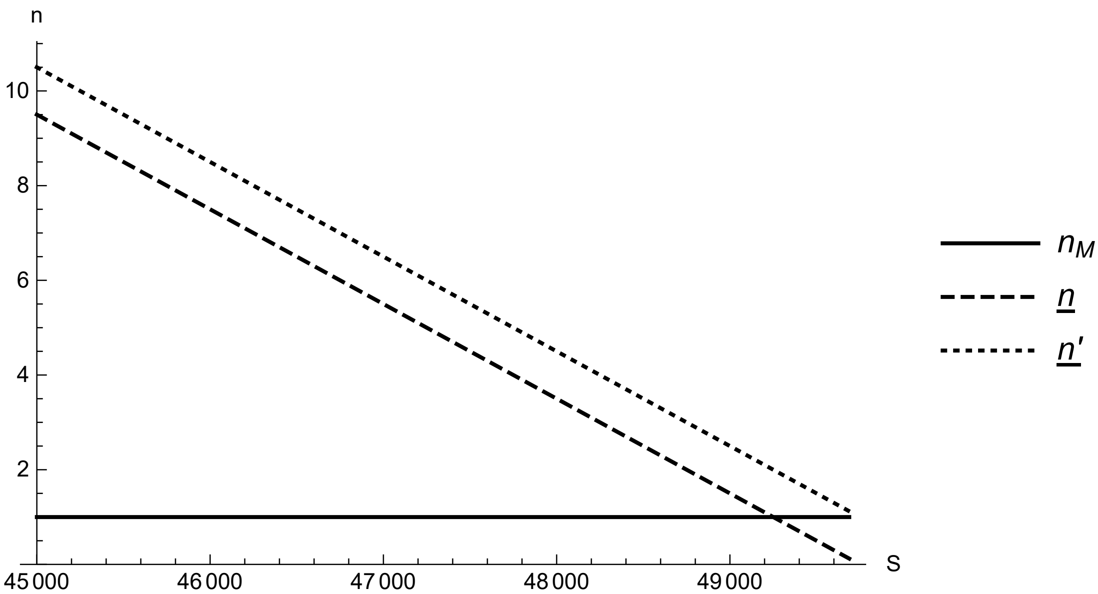


Figure 2.2.: Monopolistic and deterrent number of outlets.

3. Retail Markets

3.1. Shopping Hours and Entry - an Empirical Analysis of Aldi's Opening Hours

Authors: Samuel de Haas, Daniel Herold and Jan Thomas Schäfer

Published in: *Journal of Industry, Competition & Trade*, vol. 20, 2020, pp. 139-156.

Journal of Industry, Competition & Trade (2020) 20:139–156
https://doi.org/10.1007/s10842-019-00316-1

Shopping Hours and Entry - an Empirical Analysis of Aldi's Opening Hours

Samuel de Haas[1] · Daniel Herold[1] · Jan T. Schäfer[1]

Received: 17 April 2019 / Revised: 1 July 2019 /
Accepted: 7 August 2019 / Published online: 27 August 2019

Abstract

Aldi, the biggest discounter in Germany, started to systematically extend shopping hours of its stores in 2016. We interpret the decision to extend opening hours of a specific Aldi store as entry into a new market. By using a novel data set containing the opening hours of nearly all German grocery retailers, we find the following interesting correlations: The probability that a given Aldi outlet extends its shopping hours past 8 p.m. (i) increases if nearby Aldi outlets already extended shopping hours and (ii) decreases if nearby stores run by Aldi's close competitors did not expand shopping hours past 8 p.m.. These results seem surprising in conjunction with cannibalization and residual demand, but can be explained by consumer and firm learning or market expansion.

Keywords Shopping hours · Retailing · Coordination · Market entry

JEL Classification L22 · L41 · L81

1 Introduction

In mid 2015, Aldi, Germany's biggest discounter,[1] started to extend the opening times of some of its stores from 8 to 9 p.m.. This article empirically examines the drivers of the decision to open stores longer. For this purpose, we collected data on the opening hours of all German retail grocery stores as well as the distance between them. By controlling for socio-demographic variables such as income and population density, we study the strategic effects driving the decision of Aldi to extend shopping time of a given store until 9 p.m..

[1] Aldi has 4,195 stores and Lidl, the second largest discounter, 3,187 outlets (See Section 5).

We would like to thank the anonymous referees as well as Niklas Dürr, Justus Haucap, Paul Heidhues, Michael Hellwig, Georg Götz and Peter Winker for discussing earlier versions of this article.

✉ Samuel de Haas
samuel.de-haas@wirtschaft.uni-giessen.de

1 Justus-Liebig-University, Chair for Industrial Organization, Regulation and Antitrust, Licher Strasse 62, 35394 Giessen, Germany

The goal of our study is to examine how the choice to open a given Aldi outlet longer than 8 p.m. is affected by the opening hours of nearby stores run by competitors as well as by nearby Aldi outlets. A quasi-natural experiment in the industry allows us to isolate these strategic responses of a given Aldi store because until mid-2015 basically all Aldi outlets were closing at or before 8 p.m.[2] This schedule was a world-wide policy: By the time we collected the data for this study, for instance, US shops were also closing at 8 p.m.[3]

Although speculative, we assert that this change in policy is due to a turnover in Aldi's leadership structure.[4] Despite being unable to pin down the exact reasons for these developments, Aldi de facto refrained from extending shopping hours after 8 p.m. for a long period of time after legal changes in 2006 and 2007 allowed the extension of shopping hours. All of Aldi's main competitors indeed extended opening hours after these changes in regulations came into force (see Section 4.2). We thus assume that around 2015 there was an exogenous event that triggered Aldi's strategic expansion of opening hours.

In 2015, four grocery retail groups controlled roughly 67 % of the German market: Edeka (25.3 %), Rewe (15 %), Schwarz (Lidl and others) (14.7 %), and Aldi (11.9 %). There were 6 smaller players with 1–5 % market share whereas there is a fringe of 13.7 % that comprises very small suppliers.[5] Since at least 2007, most federal states of Germany allow retail grocery stores to open longer than 8 p.m. on workdays.[6]

To examine the drivers of Aldi's decision on whether to extend the shopping hours of a given store, we interpret this decision as entry into a new market, namely, the market of retail grocery shopping after 8 p.m. This interpretation allows us to apply a straightforward Logit regression in order to analyze entry decisions based on, e.g., Berry (1992) and Toivanen and Waterson (2005) (see Section 3). To the best of our knowledge, we are the first to adopt this approach to examine shopping hours. Moreover, studies thoroughly examining the latter topic empirically are quite scarce (see Section 2).

Our main findings are, first, that it seems that firm learning influences Aldi's decision to expand shopping hours of its outlets. In our sample, the probability that a given Aldi outlet extends shopping hours past 8 p.m. decreases with the number of outlets run by Aldi's close competitors closing at or before 8 p.m.. However, the number of competitor's outlets that are located close to a given outlet of Aldi and that allow for shopping past 8 p.m. does not significantly influence Aldi's decision about whether to extend shopping hours of that outlet. This observation suggests that Aldi learns from its closest competitors in which areas not to open later than 8 p.m. Second, we find that it seems that consumer learning occurs. The probability that a given Aldi outlet is open until 9 p.m. significantly increases in the number of nearby Aldi outlets that are also open until 9 p.m. Controlling for population density and purchasing power, this result thus provides evidence that consumers learn in which areas shopping at Aldi is possible until 9 p.m., which makes it profitable for other Aldi outlets located close to these areas to also extend shopping.

The findings about consumer learning can be interpreted from the viewpoint of evolutionary game theory as proposed by Kosfeld (2002). In those areas where coordination

[2] Exceptions are some markets located in shopping centers. Systematic extensions of opening hours started in 2015, see https://goo.gl/aRG5fs

[3] See, https://goo.gl/gWFY0U or Aldi's online store locator (https://goo.gl/50OpWj) for detailed information.

[4] Aldi is and always was a family-led firm and information on its policies and plans are extremely scarce. After the death of the two founders in 2010 and 2014, respectively, changes in the firm's conduct were expected by experts (See https://goo.gl/V2kycS; https://goo.gl/h7Mz0H; https://goo.gl/Y9yJDc).

[5] See BVE (2016); https://goo.gl/TibsZv.

[6] Exceptions are Bavaria and the Saarland where stores are allowed to be open from 6 a.m. until 8 p.m. only.

between consumers and discounters was successful, extending shopping hours of an additional Aldi store is more likely to be profitable. This pattern may also be explained by market expansion, where consumers demand increases when shops they are loyal to expand their shopping hours (Yamada (2019), Flores and Wenzel (2016)). Moreover, one can interpret the aforementioned firm learning in a similar fashion. In areas where the population of consumers focused on discounters have not (yet) coordinated on longer shopping hours, Aldi refrains from opening longer than 8 p.m.. These may be those areas where close competitors also close their outlets at or before 8 p.m..

The article is structured as follows. In Section 2, we outline the determinants of shopping hours as well as the related literature. In Section 3, the model upon which our empirical approach is based will be explained. Some facts about the German retail grocery industry are outlined in Section 4. In Section 5, we present our empirical results as well as descriptive statistics. Section 6 concludes and robustness checks are stated in the Appendix.

2 Related Literature

Retail grocery stores compete in different dimensions. First, they compete in prices (see, e.g., Binkley and Connor (1998)). Second, the quality of products and services may differ between, for instance, full-range retailers and discounters (Ellickson 2006). Third, retailers may strategically locate their stores in order to differentiate from competitors (Ellickson and Grieco 2013; Konishi 2005). Moreover, they compete in at least a fourth dimension, namely, in their shopping hours.

Most of the existing theoretical literature on shopping hours aims at determining welfare-maximizing opening hours (see, e.g., Thum and Weichenrieder (1999) and Shy and Stenbacka (2006) for an overview). Some authors (Clemenz 1990; Ferris 1991; De Meza 1984) argue that opening hours constitute another instrument of competition retailers can use in order to differentiate from competitors and thus to raise prices. One major factor potentially decreasing social welfare is that retail prices may increase in the aftermath of an extension of shopping hours, which can be interpreted in the light of increased product differentiation.

According to Shy and Stenbacka (2006), there are four factors influencing opening hours of stores. These are:

1. Spillover effects: Consumers choose shopping hours other than their preferred ones when a given set of stores does not offer their services at the consumers' most preferred shopping time.
2. Asymmetric ideal service times: Consumers' schedules lead to peaks in demand on a regular basis, e.g., on Saturday morning when many consumers have the time to get their weekly grocery supplies.
3. Economies of Scale: The cost structure of the store may exhibit increasing, constant or decreasing returns to scale with respect to sales and/or service hours.[7]
4. Price competition: As a reaction to intense price competition, suppliers may change their opening hours. Firms thus use opening hours as a means to relax competition and increase prices.

[7]Note that it is possible that the cost function is not continuous because of, e.g., legal restrictions on the maximum number of working hours per employee.

Shy and Stenbacka (2006) analyze optimal opening and closing times based on a Salop model.[8] By allowing for non-uniformly distributed consumers' tastes on shopping hours, they show that spillovers can occur whereby consumers adjust their shopping time. This creates social (transportation) costs and Shy and Stenbacka (2006) show that a monopolist provides a time frame for shopping that is narrower than the one that would be welfare maximizing. By including a competitor into their model, they show that, in equilibrium, there may be a first-mover advantage that allows one firm to offer non-stop shopping (i.e., providing the good over the whole circle). The second mover provides its services only for a limited amount of time, i.e., it covers only a smaller part of the circle.

Inderst and Irmen (2005) analyze price competition (see point 4 above) in a duopolistic, two dimensional Hotelling model where suppliers offer services differentiated in a spatial and a temporal dimension. One result of their model is that deregulating shopping hours allow for differentiation on the time dimension resulting in relaxed competition between the competitors. Hence, both stores will charge higher prices in equilibrium. Interestingly, in the mid-1990s, some German consumers anticipated an increase in retail prices when shopping hours were deregulated (Thum and Weichenrieder 1999).

Yamada (2019) extends the analysis of Inderst and Irmen (2005) by introducing quality investments into their setup. She finds that liberalization yields an equilibrium with maximal differentiation in business hours. That is, one retailer opens for a longer and another opens for a shorter period of time in equilibrium. She also finds that under liberalization, a retailer with longer (shorter) business hours charges higher (lower) prices.

The result that asymmetric shopping hours can arise in equilibrium is particularly interesting in terms of our data set because we observe a pattern of differentiated shopping hours. However, the structure of the natural experiment we utilize does not allow for testing the results of Shy and Stenbacka (2006), Inderst and Irmen (2005), or Yamada (2019) because data from the time of the liberalization of shopping hours (e.g., from 2004–2008) are required in order to do so. Moreover, testing these theories requires data on retail grocery prices or price indices.

Based on evolutionary game theory, Kosfeld (2002) analyzes the empirically observed phenomenon that German retailers reduced their shopping hours again after extending them shortly after it was possible from a legal perspective. After the first deregulation of shopping hours in Germany in 1996, about 60 % of the shops expanded their opening hours. However, 1 year later, some 10 % points of total retailers reduced the opening hours again. Kosfeld (2002) interprets this observation as a result of the consumers' and retailers' failure to coordinate on a potentially efficient equilibrium. It is costly for retailers to expand shopping hours and at the same time consumers face uncertainty: if they decide to shop later, they will not know for sure whether the store is indeed open. The resulting coordination problem did potentially render an expansion of shopping hours unprofitable in some regions.

Our findings can be interpreted against this background. We find that the probability that a given Aldi outlet expands shopping hours increases when there is an Aldi store nearby that is active in the market after 8 p.m.. This suggest that the consumers in this area have learned that shopping until 9 p.m. is possible at Aldi stores. Although Toivanen and Waterson (2005) cannot identify such a pattern for the UK fast food market, our results are consistent with consumer learning or habit formation in a similar fashion as in Kosfeld (2002) having influenced the decision of some Aldi outlets whether to allow for shopping later than 8 p.m..

[8] Other models analyzing spatial and temporal competition are Hosseinipour and Sandoh (2013) and Sandoh et al. (2015).

Related to our analysis is the concept of market expansion (Yamada (2019), Flores and Wenzel (2016)). In the context of retailing, market expansion refers to the effect that a group of consumers loyal to a certain retailer may expand their demand if shopping hours are extended. Thus, in contrast to models where total demand is not affected by business hours (e.g., Shy and Stenbacka (2006)), the expansion of shopping hours may have a positive effect on overall demand. Flores and Wenzel (2016) find that this market expansion effect of longer opening hours potentially lead to price increases, which can be detrimental to welfare. Anticipating our results, the finding that the probability that a given Aldi outlet expands its shopping hours increases in the number of nearby Aldi outlets that already have expanded their shopping hours may also be explained by market expansion, i.e., Aldi tries to generate additional demand by giving loyal consumers the opportunity to shop later. Given the lack of price data, however, we cannot analyze potential changes in the price levels.

Another central ingredient to our empirical approach is entry theory. Firm entry has been the object of numerous studies.[9] For instance, theoretical models deal with entry deterrence (e.g., Dixit (1979)) or product differentiation (e.g., Shaked and Sutton (1982)). The expansion of shopping hours may be interpreted also against the background of entry deterrence in a sense of pre-emption: in order to achieve the highest possible overlap with consumer preferences, an outlet may be open as long as possible so that an expansion of rival outlets' shopping hours becomes unprofitable. Against empirical observations on entry deterrence (Smiley 1988), this constitutes a potential strategy of entry deterrence . However, we do not find evidence for entry deterrence in our data. Given that we interpret the decision to expand shopping hours as entry into a new market (see Section 3), entry of an Aldi outlet would be deterred if the number of nearby rival outlets that already expanded shopping hours did negatively influence the propensity of that Aldi outlet to expand shopping hours. Anticipating the empirical findings, we cannot identify such a pattern.

An additional aspect from a labor market perspective, which is relevant for our study, is the concept of "threshold labor" Gradus (1996), i.e., staff that should be present irrespective of the sales volume. Additional staff in the stores constitutes a cost factor of extending shopping hours. Whereas expenses such as additional cooling, heating, or light may be negligible, the labor costs of longer opening hours may be not. We qualitatively address the supply of labor as well as labor costs in different regions by including unemployment rates as covariates.

3 Theory

We interpret the decision to extend shopping hours of a given Aldi outlet as entry into the market of shopping after 8 p.m. This departure from the traditional Hotelling-based analyses (e.g., Inderst and Irmen (2005); Shy and Stenbacka (2006)) allows for a convenient and tractable empirical assessment.

Our approach is largely based on the model of Toivanen and Waterson (2005) who analyze entry into the fast-food market in the UK. As it is standard in the Industrial Organization literature, entry occurs if expected profits from entry into a new market exceed those from not entering. In our framework, this means that a given Aldi outlet's shopping hours will be

[9] A complete review of this literature is beyond the scope of this article. An early review of the empirical literature on entry can be found in Geroski (1995). We will outline the relevant links to that strand of literature in Section 3.

extended if, all else equal, the additional profits exceed the costs from extending shopping hours.[10] Formally, let $e_{ijt} \in \{0, 1\}$ denote the decision of outlet $i \in I$, where I is the set of outlets, in market j at time t on whether to enter a particular market ($e_{ijt} = 1$ and $e_{ijt} = 1$ denotes entry and no entry, respectively).[11] Profits π_{ijt} are a function of e_{ijt}, of the number of own outlets Own_{ijt} and the number of rival outlets $Rival_{kjt}$ where k is an index for the respective rival. Let $\boldsymbol{Rival}_{jt}$ denote the vector containing the rival outlet's entry decisions in market j at time t. Entry costs are denoted by F_{ijt}. Total costs of entry are furthermore influenced by an i.i.d. cost-shock ϵ_{ijt}. Similar to (Toivanen and Waterson 2005) total profits can be formalized as follows.

$$\Pi(e_{ijt}|Own_{jt-1}, \boldsymbol{Rival}_{jt-1}) = \pi(Own_{jt-1}+e_{ijt}, \boldsymbol{Rival}_{jt-1}) - (F_{ijt}(Own_{jt-1}+\epsilon_{ijt}))e_{ijt} \; \forall k \quad (1)$$

Just as in Toivanen and Waterson (2005), we model a sequential entry decision as proposed by Berry (1992). The rationale for this assumption is that Aldi started to extent shopping hours of its outlets past 8 p.m. in 2015, years after its rivals did. For our analysis, the exact order of movements is not important as long as Aldi is the last firm to decide about entry. Thus, Aldi takes the entry decisions of its rivals as given. Those entry decisions are collected in the vector $\boldsymbol{Rival}_{jt-1}$. With this approach, we avoid potential problems arising from the existence of multiple equilibria. The existence of multiple equilibria may render the economic interpretation of the results ambiguous (see Ellickson and Misra (2011) for an overview). For the assumption of sequentiality to hold in our model, it is essential that rival outlets did not strategically react to an Aldi outlet's decision whether to extend shopping hours. In our data set, we do not find evidence of such reactions (see Section 5.2).

Central to the approach of Toivanen and Waterson (2005) is the idea of firm learning. If there is an outlet of a competitor in the industry, this not only implies that residual demand for the firm under consideration is smaller in expectation, but it also signals that the competitor runs a profitable business. When demand is uncertain, entry can thus be fostered by a competitor's activity. The entrant learns from the competitor's mere presence in the market that entry is profitable. The source of uncertainty we consider relevant is the (unobservable) consumer preference to shop later than 8 p.m. Consequently, the residual demand for a given Aldi outlet is unknown in our setup.

For the decision whether to extend the opening hours of a given Aldi outlet, the entry decisions of all the competitors' outlets are exogenously given at the beginning of each period t. In particular, the number of the competitors' stores in the relevant distance around an Aldi outlet being opened or closed after 8 p.m. is exogenous variables (see above). We thus assume that Aldi may draw information from the entry decisions of its competitors in the respective markets. Note that this also applies to Aldi stores which extended their opening times before period t. This assumption will be further discussed in Section 5.

Although we have not access to respective data, we expect that entry costs are relatively small in our example. Just as Toivanen and Waterson (2005), we are unable to distinguish between fixed and sunk costs of entry. However, besides some very minor costs such as

[10]This framework ignores potential effects of cannibalization within the Aldi group, i.e., between different Aldi outlets. Anticipating the empirical findings, cannibalization seems to be no issue in our empirical analyses because an Aldi outlet's probability to extend shopping hours increases in the number of nearby Aldi outlets where shopping hours have already been expanded.

[11]Note that contrary to Toivanen and Waterson (2005), we do not consider a duopoly. In the baseline model, we consider the outlets of all four major German food retailers as explained in the introduction.

changing the websites or advertising brochures, the main cost driver of an extension of opening hours is most likely the threshold labor value (Gradus 1996), i.e., providing the staff necessary to offer an additional hour of shopping. Although labor costs are typically seen as variable cost, in our case they are fixed because each period in the market entails these labor costs, irrespective of the number of units sold in a given outlet.[12]

Another point worth mentioning is the issue of (self-)cannibalization at a given outlet. By extending shopping hours, some consumers of a given Aldi outlet might switch from shopping before to after 8 p.m. These consumers, however, still buy at the given Aldi outlet. If, on the contrary, Aldi did not allow for longer shopping hours, some consumers might instead switch to a competitor. A positive correlation between the number of a nearby rivals outlets where shopping after 8 p.m. is possible and the probability that a given outlet is open after 8 p.m. could thus also be explained as an attempt to prevent business stealing by competitors or by Aldi itself.

4 The German Supermarket Industry

4.1 General Characterization of the Industry

The German supermarket industry is characterized by a high level of concentration. The largest share of the market is divided among the Edeka Group, the Rewe Group, the Schwarz Group, Aldi, and the Metro Group. With the takeover of Kaiser's Tengelmann by Edeka in 2015, it can be assumed that the share of the top 5 will increase further.[13]

Edeka Group, a cooperation of independent goods retailers, is the largest operator of supermarkets in Germany. Several brands belong to the group's portfolio. "Netto" covers the discount segment and Edeka markets are either active in the supermarket or the hypermarket segment. Similar to Edeka Group, the Rewe Group was founded as a cooperation of independent retailers. The largest subsidiaries are "Rewe" (supermarkets and hypermarkets) and "Penny" (discounter). The Schwarz Group operates discount markets ("Lidl") as well as hypermarkets ("Kaufland"). Aldi consists of two independent units, Aldi Nord and Aldi Sued and solely operates discount markets. It is said to be one of the price-leading companies in the market (Morschett et al. 2006).

Aldi and Lidl as well as other discount chains are typically characterized by a limited selection of products and a scarce supply of branded products. The sales area as well as the equipment are typically functional, with less services being offered. The discounters Aldi and Lidl claim that their own brands are of equal quality as those of leading brands but are cheaper. Often, products are in fact produced by or in cooperation with those brands; however, the actual manufacturer is often veiled. Due to the growing intensity of competition, traditional supermarkets like Edeka and Rewe have also started to add generic products to their assortment. The prices of those products are close to those at Aldi and Lidl.

Supermarkets offer a large variety of branded products. The sales area is larger and there are more services offered than in discount markets. However, Aldi and Lidl have started to offer additional branded products and to increase competitive pressure on competitors and also among each other over the past years.[14] Moreover, the quality of the sales areas is to be

[12]Typically, more employees are present during daytime. However, also at nighttime at least some employees have to be present if the outlet is open.

[13]See BVE (2016); https://goo.gl/TibsZv.

[14]See Lebensmittelzeitung, 25, 19.06.2015, https://goo.gl/8FRPST

improved to match those of supermarkets such as Edeka or Rewe, and Aldi and Lidl plan to offer a wider range of services like rest rooms.[15] As prices and quality are becoming more and more equal, opening hours constitute one of the only remaining strategic variables among which Aldi was not aligned with its competitors before 2016.

While store managers of supermarket chains such as Rewe and Edeka are allowed to adjust strategic variables to some degree, within the Aldi group the decision-making process is completely centralized. Aldi applies nationwide uniform pricing, i.e., there are no differences in prices between different Aldi outlets across Germany. The size of stores and variety of products clearly vary within the Edeka Group and Rewe Group, whereas the stores of Aldi and Lidl are largely standardized. Even the arrangement of the products is very similar for Aldi stores all over Germany (Brandes and Brandes 2015). We exploit this homogeneity between Aldi stores in our empirical analysis: we can rule out that the decision to extend shopping hours of a specific Aldi store is driven by internal factors like size of the shop or product variety. Instead, the decision seems to be driven by external factors like socio-demographic variables or the competitive environment in the local market, respectively.

4.2 Regulation of Shopping Hours in Germany

Since 2006, the federal states of Germany are responsible for the regulation of the opening hours within the respective state. Since 2006 and 2007, shopping hours are allowed to be extended past 8 p.m. in Baden-Wuerttemberg, Berlin, Brandenburg, Hamburg, Hesse, Lower Saxony, and Schleswig-Holstein. In Mecklenburg-Western Pomerania, North Rhine-Westphalia, Saxony-Anhalt, Thuringia, Rhineland-Palatinate, and Saxony regulations vary according to the day of the week. Only in Bavaria as well as the Saarland, no changes have been made, and opening hours are regulated the same way they were before 2006 (i.e., 8 p.m. as an upper limit). On Sundays, the operation of shops is only possible as an exception and strict conditions have to be met in order to be open, regardless of the federal state.[16] As discussed above, Edeka, Rewe, and Lidl started to extend shopping hours of their shops after the deregulation. Aldi, however, refrained from opening its stores longer than 8 p.m. until 2015. In 2016, they started to systematically extend shopping hours of various stores—this observation constitutes the starting point for our analysis.

5 Empirical Analysis

In this section, we present our empirical analysis. We first describe the construction of our data set in Section 5.1 for which we derive descriptive statistics including information on number, locations, and opening hours of stores in Section 5.2. In Section 5.3, we present the results of our regressions which estimate the impact of different socio-demographic variables and the opening hours of nearby competitors on the decision to extend shopping hours of a specific Aldi store.

[15] See Lebensmittelzeitung, 40, 02.10.2015, https://goo.gl/pcToCL and ALDI Press Release, 11.05.2016, https://goo.gl/MIP4yP.

[16] See law on shop opening hours of each federal state (Ladenöffnungsgesetz) as well as the federal law on business hours (Gesetz über den Ladenschluss).

5.1 Construction of the Data Set

Our data set was constructed upon merging different sources. The primary data were obtained by collecting shop data including addresses and opening hours of all Aldi, Lidl, Rewe, Edeka, Penny, and Netto stores in Germany listed in the corresponding online store finders.[17] We collected these data for two points in time: July ($t = 0$) and December ($t = 1$) 2016. As a next step, we compute distances—more specifically, the driving time—of every single Aldi store to all its competitors and other Aldi stores. We restrict our attention to those stores located within a maximum range of 20-min driving time. In doing so, we implicitly define local markets and control for the competitive environment within these markets, i.e., the number of nearby competitors for all Aldi stores.[18] The driving times are based on Open Source Routing Machine using the fastest route option.[19] As control variables, we include the following socio-demographic variables: population density, private purchasing power per household, the unemployment rate, the share of households under 40 years of age (based on the head of the household), the share of 1-person households, and the share of 2-person households in our regressions.[20] The first three variables are well-known covariates in microeconometric analyses. The last three parameters (households under 40 years of age, share of 1-person households, and share of 2-person households) are added to the analysis because we assume younger people and one- or two-person households to have specific preferences for shopping hours based on their working habits. For instance, a one-person household may be more inclined to shop after work in the evening whereas in a household with children shopping trips may take longer so that the preferred shopping time may be on Saturday morning.

There are two potential caveats in using those socio-demographic covariates. First, the data is available on postcode level only. Consequently, spatial separation on the demand side is not as detailed as the aforementioned radius around the respective Aldi outlets. Second, the data is only available on an annual basis whereas we observe variations during the year. However, we expect the impact on the estimation to be negligible because we use these data only as controls for socio-demographic effects.

5.2 Descriptive Analysis

Our data set contains roughly 23,000 stores. Table 1 provides an overview of the suppliers' market shares as well as the share of stores with opening hours later than 8 p.m. at $t = 0$ (July 2016).

We find that, with about 77%, Rewe has the highest share of stores that are opened after 8 p.m. The share is even higher (91%) if we exclude those stores in Bavaria and Saarland where shopping later than 8 p.m. is generally not allowed (see Section 4.2). Edeka has a share of about 27% (36% without Bavaria and Saarland) of stores that are opened after 8 p.m. The shares of the discounters are 72% (86%) for Lidl, 61% (70%) for Penny, and 43% (53%) for Netto. With about 23% (26%), Aldi has the smallest share of stores with long

[17] See https://filialfinder.aldi-sued.de/, https://www.lidl.de/de/filialsuche/, https://www.edeka.de/marktsuche.jsp, https://marktsuche.rewe.de/ and https://www.rewe-dortmund.de/marktsuche/ (last accessed on 30 June, 2017)

[18] For instance, the German Federal Cartel Office uses a range of 20-min driving time in order to define a local market in their analyses of the German retail grocery industry (Bundeskartellamt 2015).

[19] http://project-osrm.org/ (last accessed on 30 June 2017)

[20] This data is provided by Axciom Deutschland GmbH, Neu-Isenburg, 2016.

Table 1 Absolute number and share of long opened stores (July 2016)

Competitor	Number of stores (market share)	Opened after 8 p.m. (share)
Aldi	4195 (18.2%)	948 (22.6%)
Lidl	3187 (13.8%)	2306 (72.4%)
Rewe	3461 (15%)	2650 (76.6%)
Edeka	5920 (25.7%)	1607 (27.2%)
Penny	2133 (9.3%)	1312 (61.5%)
Netto	4116 (17.9%)	1779 (43.2%)
Sum	23,012 (100%)	10,602 (46.07%)

opening hours. These figures are not surprising because Aldi started to systematically extend opening hours in 2016 whereas the other retailers started to do so after the deregulation in 2006. However, the number of Aldi stores that were open until 9 p.m. in $t = 0$ was already higher than the initially targeted number of 650 stores.[21] Thus, it seems that the announced test phase was successfully concluded prior to our observation period and Aldi started to systematically extend shopping hours of further stores. Focusing on Aldi outlets that were closed after 8 p.m. at $t = 0$ and without stores located in Bavaria and Saarland, we find 2646 Aldi stores that could potentially enter the “new market” for shopping after 8 p.m..

Aldi has the largest number of shops extending their shopping hours during our observation period (from $t = 0$ to $t = 1$, i.e., from July to December 2016). They extended opening hours of 85 stores. Lidl extend opening hours of 4, Rewe of 10, Edeka of 13, Penny of 7, and Netto of 32 stores. These numbers are relatively small so that we do not expect them to be the result of strategic considerations concerning the changes in Aldi’s policy. More importantly, even if these adjustments in the rivals shopping hours would have been reactions to Aldi’s policy, there would be no impact on our estimation results. This is because none of the adjustments in the respective stores’ shopping hours occurred within the relevant radius (20-min driving time) around an Aldi outlet that extended its opening hours during our observation period. There was also a handful of newly opened and closed stores for all competitors. However, we are not interested in these numbers because we assume that the decision to extend opening hours of a specific Aldi store in $t = 1$ is based on the competitive environment in $t = 0$ ($Rival_{k,j,t-1}$), as described in Section 3.

Figure 1 illustrates the locations of the corresponding stores all over Germany by ownership and opening hours in $t = 0$. It becomes obvious that Rewe outlets stayed open after 8 p.m. almost everywhere, except for Bavaria and Saarland, where opening hours are still restricted to 8 p.m. There are also some stores with shorter opening hours in North Rhine-Westphalia. These outlets are mostly owned by “Rewe Dortmund” which is the only independent regional subsidiary of Rewe. Most of Lidl’s outlets in Western Germany and Berlin offer shopping hours past 8 p.m. whereas only a few Lidl stores in Eastern Germany are open after 8 p.m. A similar pattern appears when examining Penny’s and Netto’s outlets. For Edeka, a somewhat different picture emerges. The share of Edeka outlets with closing times at or before 8 p.m. in Western Germany, especially in the northern parts, is much higher than the corresponding shares of Lidl, Penny, or Netto. For Aldi, we find that stores with opening hours later than 8 p.m. are mostly located in big cities and urban areas with high population densities such as Hamburg, Berlin, or the Rhine-Maine and

[21] See Handelsblatt online, https://goo.gl/Ke9NTx (last accessed on November 7, 2016)

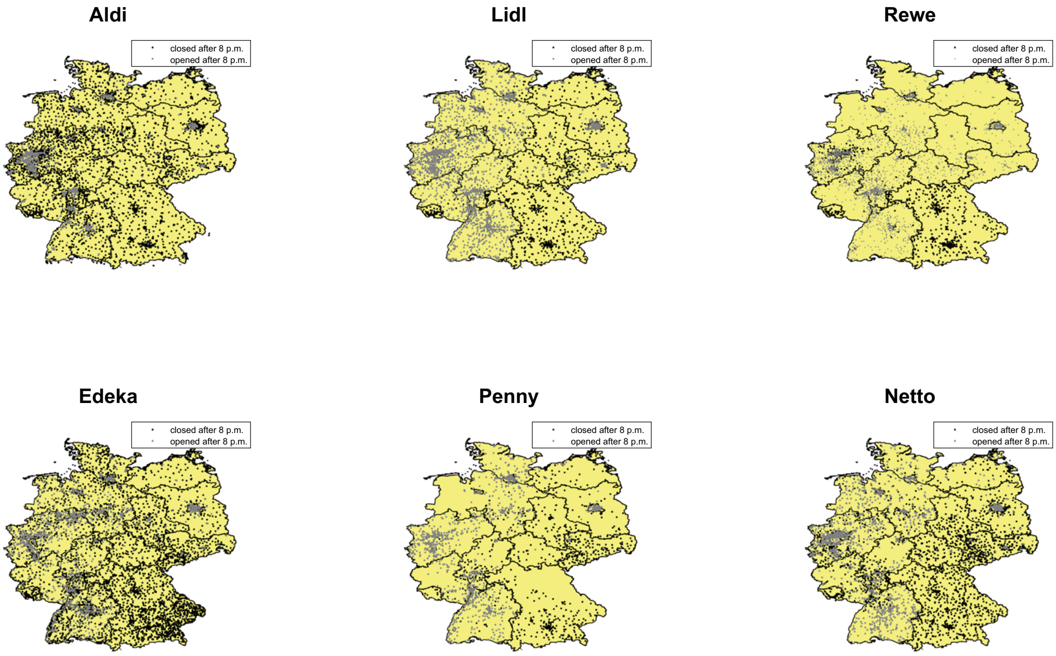


Fig. 1 Location and shopping hours of the stores in July 2016

Ruhr area. The share of Aldi stores that are open after 8 p.m. is particularly high in North Rhine-Westphalia.[22]

We observe that Rewe follows a country-wide strategy that does not seem to be primarily influenced by those strategic or even socio-demographic factors we consider in our analysis. We therefore interpret Rewe as the first mover or the incumbent in the market for shopping after 8 p.m. Lidl's, Penny's, and Netto's strategies can be characterized by a high degree of differentiation between Eastern and Western Germany. Although there was an area-wide coverage of shops being opened later than 8 p.m. in Western Germany, outlets in Eastern Germany had shopping hours later than 8 p.m. mostly in regions with a high population density. For Edeka, there does not emerge a clear pattern. One explanation for this observation might be that the owners of an Edeka outlet are more independent than those of the competitors and may be able to adjust their shopping hours more freely.[23]

5.3 Empirical Results

As described in Section 3, we assume that the decision to extend shopping hours of a specific Aldi store at time t (December 2016) is based on its competitive environment at $t-1$ (July 2016). Thus, we estimate the following Logit model (Toivanen and Waterson 2005):

$$Pr(e_{j,t}|X_{j,t-1}, Z_{j,t-1}) = \delta_0 + X'_{j,t-1}\beta_1 + Z'_{j,t-1}\beta_2 + \epsilon_{j,t} \quad (2)$$

By using this Logit model, we estimate the probability that a given Aldi outlet extend its shopping hours. More specifically, we estimate the impact of the explanatory variables on

[22]We conjecture that this observation can be explained by the close distance to Aldi's headquarters located in Essen and Mühlheim, where they potentially started to extend shopping hours of its stores.

[23]Most Edeka outlets are owned by independent retailers. See also Edeka concept http://goo.gl/rtUaEf (last accessed on December 8, 2016).

Table 2 Logit estimation results

	Logistic regression		Marginal effects	
$AldiOpened$	32.016**	(15.821)	0.9673*	(0.4944)
$LidlOpened$	−0.09914	(1.1271)	−0.002995	(0.03421)
$LidlClosed$	−115.99*	(65.680)	−3.5045*	(2.0454)
$ReweOpened$	−0.08515	(0.5685)	−0.002573	(0.01720)
$ReweClosed$	−34.249	(33.452)	−1.0347	(1.0423)
$EdekaOpened$	0.1479	(0.1705)	0.004468	(0.005235)
$EdekaClosed$	−0.7993	(1.3204)	−0.02415	(0.04042)
$PennyOpened$	−5.2418	(7.7536)	−0.1584	(0.2398)
$PennyClosed$	−74.476	(54.165)	−2.2501	(1.7330)
$NettoOpened$	−1.6887	(3.4765)	−0.05102	(0.1037)
$NettoClosed$	−29.474**	(14.269)	−0.8905**	(0.4427)
$Population$	−0.00001050	(0.00004678)	−3.172e-07	(0.000001418)
$Household < 40years$	10.454**	(4.5172)	0.3158**	(0.1371)
$PurchasingPower$	0.0001277***	(0.00004403)	0.000003859**	(0.000001591)
$Unemployment$	19.065**	(8.8790)	0.5760*	(0.2980)
$1Person - Household$	14.697***	(4.8365)	0.4440***	(0.1720)
$2Person - Household$	21.102***	(7.5135)	0.6376***	(0.2384)
Constant	−26.099***	(5.4005)		
Observations	2547			
Pseudo-R^2	0.1396			

The estimation is performed by using a logistic regression. Cluster-robust standard errors (clustered on district level) are presented in parentheses. Marginal effects after logit are obtained by using Stata's 'margins' command which treats all factor variables as balanced. Statistics are significant for $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

the probability that a given Aldi outlet extends its shopping hours. Thereby, the left-hand side takes the value 1 if Aldi outlet j extended shopping hours during our observation period and is 0 otherwise. The vector X includes the weighted[24] number of nearby (max. 20-min driving time to Aldi j) Aldi stores already opened later than 8 p.m. as well as the weighted number of nearby Lidl, Rewe, Edeka, Penny, and Netto stores closing late (past 8 p.m.) and early (at or before 8 p.m.) in $t-1$. All socio-demographic covariates are collected in the vector Z: population density, the share of households under 40 years of age based on the head of household, purchasing power per household (in Euro), share of one- and two-person households, and unemployment rate in the same postcode area as Aldi j.

The Results of the Logit regression are shown in Table 2. The terms "Opened" or "Closed" refer to the (weighted) number of nearby stores of the corresponding competitors. We observe that $AldiOpened$ has a significantly positive coefficient whereas the coefficients of $LidlClosed$ and $NettoClosed$ are (weakly) significant and negative. The other coefficients concerning the competitive environment are insignificant. Thus, all else equal, it is more likely that a specific Aldi outlet extends shopping hours if the number of nearby

[24] A store is weighted by inverse distances (in seconds of driving time).

Aldi stores with a longer opening time is higher. Furthermore, a given Aldi store's probability to be open after 8 p.m. decreases in the number of nearby Lidl or Netto stores with shorter opening hours. These correlations cannot be explained by traditional entry theory (e.g., Shaked and Sutton (1990)): On the one hand, a high number of own stores should have a negative impact on residual demand (cannibalization) and thus on entry. On the other hand, a higher number of Lidl or Netto stores with closing times at or before 8 p.m. should have a positive impact on demand because one would expect a certain share of (marginal) consumers to switch from a given Lidl or Netto outlet to an Aldi store which just started to provide shopping after 8 p.m.

Note that we only have two observation dates. Hence, our "panel" has a cross-sectional character, what in general makes it hard to exclude alternative explanations of our results. We control for that by including several local characteristics like population density or the share of households under 40 years. Additionally, we clustered the robust standard errors of our regressions on district levels. However, there might be still some unobserved heterogeneity that drives our results. Although the data set does not allow for exact economic interpretation of the results, the correlations are nevertheless insightful and very interesting.

Our findings may be explained by consumer or firm learning, and by a market expansion effect. This is for the following reasons.

First, habit formation could have a positive effect on the propensity that a given Aldi outlet expands business hours: Consumers learn that shopping longer than 8 p.m. is possible in areas with a high number of late closing Aldi stores ($AldiOpened$). Consequently, some consumers start shopping after 8 p.m. and the demand in this "market" increases. This implies that whenever Aldi's own outlets are closing late in the area around outlet j ($AldiOpened$), information about demand at later shopping hours is fairly precise from the viewpoint of Aldi. In such a case, Aldi can predict relatively accurately whether extending the shopping hours of outlet j pays off, hence the strong correlation. However, in an area with a high number of discounter outlets with shorter opening hours ($LidlClosed$ and $NettoClosed$), the process of consumer learning did not (yet) begin. Another potential explanation of the positive effect of the number of nearby Aldi outlets that have already extended shopping hours on the probability that a given Aldi outlet expands its shopping hours is a market expansion effect (see, e.g., Flores and Wenzel (2016)). As explained in Section 2, the effect of market expansion means that the overall demand of consumers may increase when business hours are extended. This effect is connected to consumer learning: if a nearby Aldi outlet has already extended shopping hours, the demand of Aldi consumers increases which may spill over to closeby Aldi outlets. As a consequence, these closeby Aldi outlets also have an incentive to expand business hours.

Second, our results indicate that in an area with a high number of discount stores closing early ($LidlClosed$ and $NettoClosed$), Aldi potentially learns that there is little demand for shopping after 8 p.m., i.e., firm learning. Such a behavior may signal Aldi where not to extend shopping hours. This argument is important because we expect Lidl and Netto to have more experience with the profitability of late-closing outlets. This is because they entered the market for late shopping earlier than Aldi. In particular, the marginal effect of Lidl closing earlier is stronger than the one of Netto. Given that Lidl is the closest competitor of Aldi, this makes sense from a managerial point of view (Brandes and Brandes 2015).[25]

[25] Aldi and Lidl have not only have similar business policies, product ranges and prices, but Lidl is also the biggest discounter after Aldi, when focusing the turn over (see Lebensmittelzeitung online, 15.03.2018, https://bit.ly/2NysIq0, last accessed on June 13, 2019).

We find that the coefficient for $Population$ is not significant whereas the coefficients of $Household < 40years$, $PurchasingPower$, $Unemployment$, $1Person - Household$, and $2Person - Household$ are significantly positive. Hence, the local share of households under 40 years, private purchasing power, unemployment rate, and the share of one and two person households raise the probability that a given Aldi outlet extends its opening hours. We expect younger people to prefer late shopping because of their working habits. The same holds for one and two person households. Another explanation would be that younger people are already used to late night shopping, whereas older people are more used to shop early given the long time of regulated business hours in Germany. The positive influence of purchasing power and unemployment rate at the same time seems surprising. One would expect that these variables are negatively correlated because higher unemployment rates are typically connected to lower income and, consequently, lower purchasing powers. It is insightful to focus on each variable separately. On the one hand, one could argue that purchasing power is correlated with longer working times and thus with higher demand for late shopping. On the other hand, one could argue that it is easier to find additional staff in areas with a high unemployment rate thereby decreasing entry costs (”threshold labor”, as discussed in Section 3). However, this surprising result, as well as the other coefficients of the covariates, may partly be driven by multicollinearity effects. A corresponding multicollinearity test is stated in the Appendix.[26]

6 Conclusions

This article examines the drivers of Aldi’s decision on whether to extend shopping hours of particular outlets past 8 p.m. in Germany. We interpret Aldi’s decision to open a given outlet until 9 p.m. as entry into a new market, namely, the market for grocery retail shopping after 8 p.m. We use a natural experiment in order to analyze this question: Aldi started to systematically extend shopping hours past 8 p.m. in 2016, roughly 9 years later than its competitors. We collected data on locations and shopping hours of all major players in the German retail grocery sector and merged these with data on socio-demographic variables such as population density and purchasing power.

Using a Logit regression based on entry theory as in, e.g., Toivanen and Waterson (2005), we find interesting correlations. First, we find that the presence of nearby Lidl or Netto outlets that are closing no later than 8 p.m. significantly decreases the probability that an Aldi outlet extends shopping hours past 8 p.m. Second, if there are Aldi outlets nearby that already close later than 8 p.m., the probability that a respective Aldi store also allows for shopping hours past 8 p.m. will increase significantly.

Potential explanations for these results are consumer or firm learning and market expansion. Habit formation suggests that consumers have learned that shopping past 8 p.m. is possible in areas with a high number of Aldi stores already closing later than 8 p.m. Kosfeld (2002). A fraction of consumers thus starts shopping after 8 p.m. and demand in this “market” increases, which may even lead to an increase in overall at the given outlet. By the same token, in an area with a high number of discounter outlets closing at or before 8 p.m., the process of consumer learning has not (yet) been initiated or it has failed.

[26]Indeed, the test indicates that there are is multicollinearity in the estimations. However, the results are qualitatively the same when skipping (subsets of) the variables that are potentially susceptible to multicollinearity, as stated in the Appendix.

The presence of Lidl or Netto stores that are closing no later than 8 p.m. potentially allows for firm learning by reducing uncertainty about demand. An area with a high number of Lidl or Netto outlets closing already at or before 8 p.m. may signal that this area is characterized by little demand for shopping after 8 p.m. Firm learning is particularly strong with respect to Aldi's closest competitors, Lidl and Netto. In particular, these competitors have better information about the profitability of late night shopping (past 8 p.m.) in a respective area because they entered the market much earlier than Aldi. Lidl's and Netto's behavior thus provides valuable information for Aldi.

Appendix

As a robustness check, we run a Logit regression as stated in Eq. 2 with a different definition of a 'local market': Instead of a maximum of 20 minutes driving time, we assume a maximum of 15 minutes driving time. Corresponding results are shown in Table 3.

The results are qualitatively the same as the ones discussed in Section 5.3.

As mentioned in Section 5.3 it is surprising that the coefficients of the variables *PurchasingPower* and *Unemployment* have the same signs. This may in part be due to multicollinearity effects. To test this, we run a Variation Inflation Factor (VIF) test. Results are stated in Table 4.

Table 3 Logit estimation results

	Logistic regression		Marginal effects	
$AldiOpened$	38.076**	(18.548)	1.1454**	(0.5800)
$LidlOpened$	−0.2435	(1.2306)	−0.007325	(0.03737)
$LidlClosed$	−131.16*	(73.498)	−3.9456*	(2.2724)
$ReweOpened$	−0.07685	(0.5658)	−0.002312	(0.01705)
$ReweClosed$	−40.205	(38.441)	−1.2094	(1.1935)
$EdekaOpened$	0.1569	(0.1758)	0.004720	(0.005381)
$EdekaClosed$	−0.7396	(1.3146)	−0.02225	(0.04007)
$PennyOpened$	−6.3524	(9.2972)	−0.1911	(0.2860)
$PennyClosed$	−81.012	(60.823)	−2.4370	(1.9162)
$NettoOpened$	−1.4495	(2.8941)	−0.04360	(0.08565)
$NettoClosed$	−24.035*	(14.014)	−0.7230*	(0.4329)
$Population$	−0.00003346	(0.00004499)	−0.000001006	(0.000001382)
$Household < 40years$	10.448**	(4.4453)	0.3143**	(0.1353)
$PurchasingPower$	0.0001285***	(0.00004376)	0.000003865**	(0.000001591)
$Unemployment$	18.874**	(8.7657)	0.5678*	(0.2944)
$1Person - Household$	14.626***	(4.8692)	0.4400**	(0.1725)
$2Person - Household$	20.983***	(7.7953)	0.6312**	(0.2465)
Constant	−26.091***	(5.5537)		
Observations	2547			
Pseudo-R^2	0.1415			

The estimation is performed by using a logistic regression. Cluster-robust standard errors (clustered on district level) are presented in parentheses. Marginal effects after logit are obtained by using Stata's 'margins' command which treats all factor variables as balanced. Statistics are significant for $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 4 VIF of regression in Table 2

Variable	VIF
$2Person - Household$	109.70
$1Person - Household$	78.18
$PurchasingPower$	65.18
$Household < 40years$	60.16
$Unemployment$	16.27
$Population$	3.26
$AldiOpened$	2.68
$NettoOpened$	1.36
$NettoClosed$	1.32
$LidlOpened$	1.27
$PennyOpened$	1.14
$EdekaClosed$	1.08
$PennyClosed$	1.06
$ReweOpened$	1.05
$EdekaOpened$	1.03
$LidlClosed$	1.02
$ReweClosed$	1.02
Mean VIF	20.40

Indeed, the VIFs of the variables *2Person-Household*, *1Person-Household*, *PurchasingPower*, *Households<40* and *Unemployment* are problematic because they are above 10 (rule of thumb, see Wooldridge (2015)). Therefore, we run the Logit regression as stated in Eq. 2 again, but only with a subset of the problematic variables. The results are stated in Table 5.

Table 5 Logit estimation results

	Logistic regression		Marginal effects	
$AldiOpened$	40.381**	(16.526)	0.6730**	(0.3165)
$LidlOpened$	−0.6134	(1.9616)	−0.01022	(0.03242)
$LidlClosed$	−101.64*	(61.300)	−1.6940**	(0.7636)
$ReweOpened$	−0.1613	(0.4941)	−0.002688	(0.008285)
$ReweClosed$	−26.368	(30.690)	−0.4394	(0.4786)
$EdekaOpened$	0.1328	(0.1514)	0.002212	(0.002660)
$EdekaClosed$	−0.8504	(1.2612)	−0.01417	(0.02128)
$PennyOpened$	−2.6608	(4.5212)	−0.04435	(0.07789)
$PennyClosed$	−54.731	(42.924)	−0.9121	(0.6169)
$NettoOpened$	−4.1672	(8.9676)	−0.06945	(0.1494)
$NettoClosed$	−29.696**	(14.836)	−0.4949**	(0.2457)
$Population$	−0.000005445	(0.00005028)	−9.074e−08	(8.438e−07)
$Household < 40years$	7.1158**	(3.1299)	0.1186*	(0.06286)
$Unemployment$	10.196*	(6.1162)	0.1699*	(0.08876)
Constant	−5.7020***	(0.8107)		
Observations	2547			
Pseudo-R^2	0.1098			

The estimation is performed by using a logistic regression. Cluster-robust standard errors (clustered on district level) are presented in parentheses. Marginal effects after logit are obtained by using Stata's "margins" command which treats all factor variables as balanced. Statistics are significant for $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 6 VIF of regression in Table 5

Variable	VIF
$Household < 40years$	7.44
$Unemployment$	7.08
$Population$	2.54
$AldiOpened$	2.53
$NettoOpened$	1.35
$NettoClosed$	1.28
$LidlOpened$	1.26
$PennyOpened$	1.14
$EdekaClosed$	1.07
$PennyClosed$	1.05
$ReweOpened$	1.05
$EdekaOpened$	1.02
$LidlClosed$	1.02
$ReweClosed$	1.02
Mean VIF	2.20

The results are qualitatively the same as the ones discussed in Section 5.3. A corresponding VIF test shows that there are no more problems with multicollinearity, as stated in Table 6.

When skipping other subsets of variables, instead of the one skipped in the underlying regression of Table 5, the results do not change (qualitatively). Corresponding regression results are available from the authors on request.

References

Berry ST (1992) Estimation of a model of entry in the airline industry. Econometrica 60(4):889–917
Binkley JK, Connor JM (1998) Grocery market pricing and the new competitive environment. J Retail 74(2):273–294
Brandes D, Brandes N (2015) Bare essentials: the Aldi success story. Linde Corporate. Linde, Wien, oCLC: 906163625
Bundeskartellamt (2015) Beschluss im Fusionskontrollverfahren Edeka/Kaiser's Tengelmann (B2-96/14). http://www.bundeskartellamt.de/SharedDocs/Entscheidung/DE/Entscheidungen/Fusionskontrolle/2015/B2-96-14.pdf?__blob=publicationFile&v=3
Clemenz G (1990) Non-sequential consumer search and the consequences of a deregulation of trading hours. Eur Econ Rev 34(7):1323–1337
De Meza D (1984) The fourth commandment: is it pareto efficient? Econ J 94(374):379–383
Dixit A (1979) A model of duopoly suggesting a theory of entry barriers. Bell J Econ 10(1):20–32
Ellickson PB (2006) Quality competition in retailing: a structural analysis. Int J Ind Organ 24(3):521–540
Ellickson PB, Grieco PL (2013) Wal-mart and the geography of grocery retailing. J Urban Econ 75:1–14
Ellickson PB, Misra S (2011) Structural workshop paper - estimating discrete games. Mark Sci 30(6):997–1010
Ferris JS (1991) On the economics of regulated early closing hours: some evidence from canada. Appl Econ 23(8):1393–1400
Flores M, Wenzel T (2016) Shopping hours and price competition with loyal consumers. The BE J Econ Analysis Policy 16(1):393–407
Geroski PA (1995) What do we know about entry? Int J Ind Organ 13(4):421–440

Gradus R (1996) The economic effects of extending shop opening hours. J Econ 64(3):247–263
Hosseinipour A, Sandoh H (2013) Optimal business hours of the newsvendor problem for retailers. Int Trans Oper Res 20(6):823–836
Inderst R, Irmen A (2005) Shopping hours and price competition. Eur Econ Rev 49:1105–1124
Konishi H (2005) Concentration of competing retail stores. J Urban Econ 58(3):488–512
Kosfeld M (2002) Why shops close again: an evolutionary perspective on the deregulation of shopping hours. Eur Econ Rev 46:51–72
Morschett D, Swoboda B, Schramm-Klein H (2006) Competitive strategies in retailing—an investigation of the applicability of porter's framework for food retailers. J Retail Consum Serv 13(4):275–287
Sandoh H, Koide T, Kiniwa J (2015) Space-time hotelling model and its application to retail competition in a duopoly. In: Proceedings of the international multiconference of engineers and computer scientists, vol 2
Shaked A, Sutton J (1982) Relaxing price competition through product differentiation. Rev Econ Stud: 3–13
Shaked A, Sutton J (1990) Multiproduct firms and market structure. RAND J Econ: 45–62
Shy O, Stenbacka R (2006) Service hours with asymmetric distributions of ideal service time. Int J Ind Organ 24(4):764–771
Smiley R (1988) Empirical evidence on strategic entry deterrence. Int J Ind Organ 6(2):167–180
Thum M, Weichenrieder AJ (1999) Ist Ladenschlußregulierung volkswirtschaftlich effizient? https://www.researchgate.net/profile/Alfons_Weichenrieder/publication/237408138_Ist_Ladenschlussregulierung_volkswirtschaftlich_effizient/links/551a95f40cf244e9a4589ce5.pdf
Toivanen O, Waterson M (2005) Market structure and entry: where's the beef. RAND J Econ 36(3):680–699
Wooldridge JM (2015) Introductory econometrics: a modern approach. Nelson Education
Yamada M (2019) Business hours, store quality, and social welfare. J Ind Compet Trade: 1–14

3.2. 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

This paper has been submitted to the *Journal of Cultural Economics*. Following peer review, we received a request for major revision and have been invited to resubmit the manuscript.

The manuscript has been presented at the following conferences:

- MaCCI Annual Conference 2019, Leipniz Centre for European Economic Research (ZEW), 2019, Mannheim, Germany.
- 5$^{\text{th}}$ Annual Conference on Competition Policy of the Bergen Center for Competition Law and Economics (BECCLE), 2019, Bergen, Norway.
- 46$^{\text{th}}$ Annual Conference of the European Association for Research in Industrial Economics (EARIE), 2019, Barcelona, Spain. Presenter: Phil-Adrian Klotz.

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

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

December 12, 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 37% 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 (see also Clement et al. (2007), Lizzeri (1999), Marvel and McCafferty (1990), and Reimers and Waldfogel (2021)).

Cowen (2008) and Gilbert (2015) address that while reducing the importance of physical stores, e-Commerce might expand overall demand for books by attracting new customers. 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) 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 analysis suggests that, on average, the closure of one physical bookstore results in a monthly decline of 744 book sales per federal state. Between 2011 and 2017, bookstore closures collectively contributed to a decrease in monthly sales of around 1 million units, accounting for approximately 37% of the overall decline in monthly 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 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. 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 on "traditional" physical retailers. There are numerous articles analyzing competition between the two retail channels, such as Brown and Goolsbee (2002), Jin and Kato (2007), Ofek et al. (2011), Gauri et al. (2021) or Couture et al. (2021). The impact of digitization on the retail landscape 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). Particularly closely related to this article are the papers of Goolsbee (2001), Prince (2007) and Goldmanis et al. (2010). Goolsbee (2001) and Prince (2007) analyze the market for computer equipment and find evidence for intensifying competition between the two channels throughout the late 1990s. Goldmanis et al. (2010) analyzes competition between the online and offline channel 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. 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, however, analyze markets where retailers compete in prices. With fixed book prices (i.e., in the absence of price competition between retailers), 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.

The article is also related 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)[1], 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.Whether such an exemption from some fundamental pillars of competition law is justified,

[1]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.

crucially depends on fixed book prices being an appropriate tool to achieve the policy goal to promote 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). In particular, Williams, 2024 finds that, on average, book sales are higher in countries with RPM, while prices are similar. He explains this findings by a positive impact of fixed book prices on the network of bookstores.

The article is structured as follows. Section 2 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 Theory

The following section provides a brief theoretical background for the empirical analyses presented in the remainder of this article. The starting point is the notion that the online and offline retail channels have different "service" features, as was already explained in the introduction. Consumers vary regarding their preferences towards these features. This is an example of (horizontal) product differentiation, which, in the literature, is usually formalized using Hotelling- oder Salop-models where consumers' preferences regarding the two sales channels are reflected in different transport costs (Balasubramanian, 1998; Bouckaert, 2000; Chu et al., 2012; Guo & Lai, 2017; Legros & Stahl, 2019).

In our empirical setting, we study the effect of a physical bookstore's market exit. In a Hotelling- or Salop-model, *ceteris paribus* such an exit has no effect on consumers buying online or patronizing a bookstore that is not closing. Net utility of those consumers who patronized a bookstore that is closing decreases due to an increase in inconvenience costs from either patronizing another bookstore or purchasing online. This can lead to a decrease in overall demand if that increase is so high that, for given prices, consumers net utility from purchasing books at another bookstore or online becomes negative.[2]

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.

[2]Such a formalization requires a model that allows for partial market coverage, i.e., a model in which not all consumers always purchase the product by assumption. To construct such a model, one would need to depart from the assumption of symmetric inconvenience costs from buying online, as it is usually used in the literature (see above). Instead, one could assume that these costs differ between consumers. An earlier version of this paper contained such a model. It is available from the authors upon request.

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[3] 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 (see Appendix A.1 for an in-depth description of the data generation process). 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 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

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

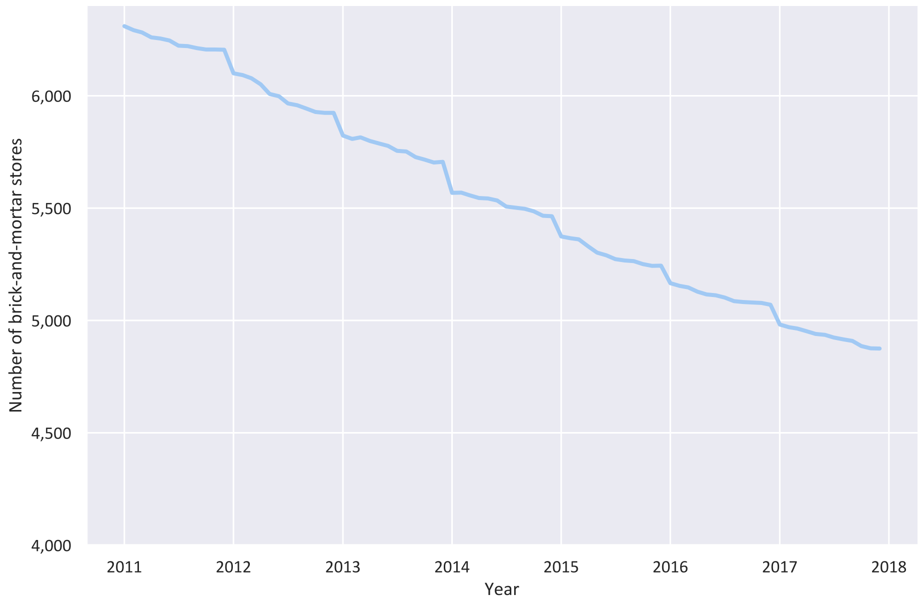


Figure 1: Development of physical bookstores over time. Source: German Publishers and Booksellers Association.

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.[4]

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.[5] Second, we use consumer panel data provided by GfK GmbH.[6] Both databases differ with respect to coverage, frequency and

[4]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.
[5]see https://www.media-control.de/buch1.html; last accessed December 20, 2023.
[6]see https://www.gfk.com/de; last accessed December 20, 2023.

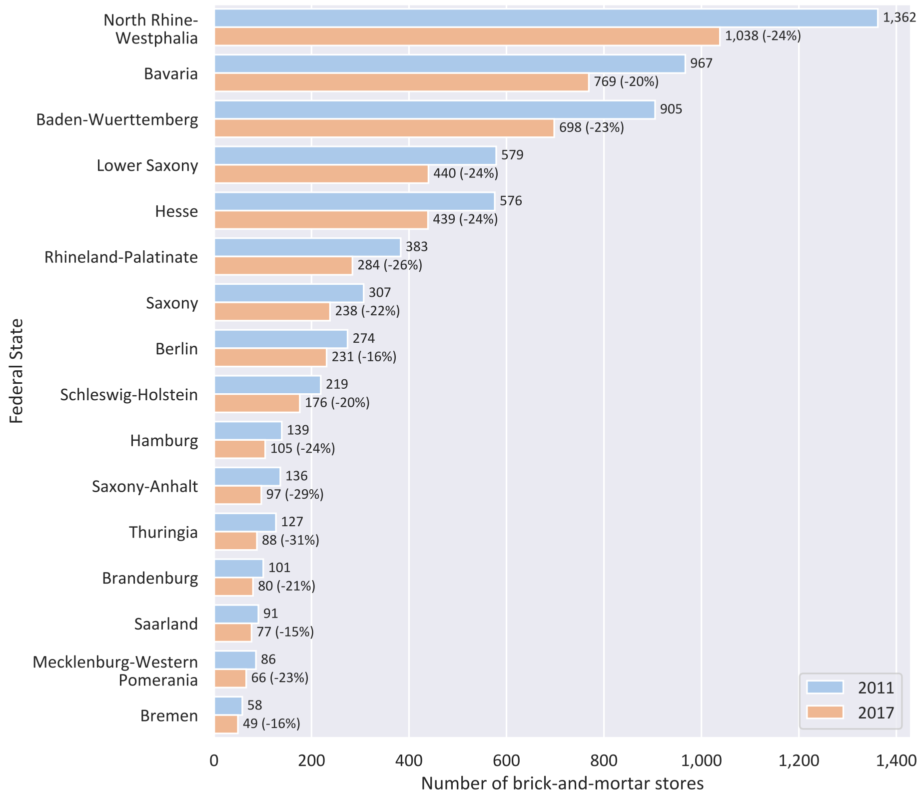


Figure 2: Development of physical bookstores on the federal state level. Source: German Publishers and Booksellers Association.

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.2 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.

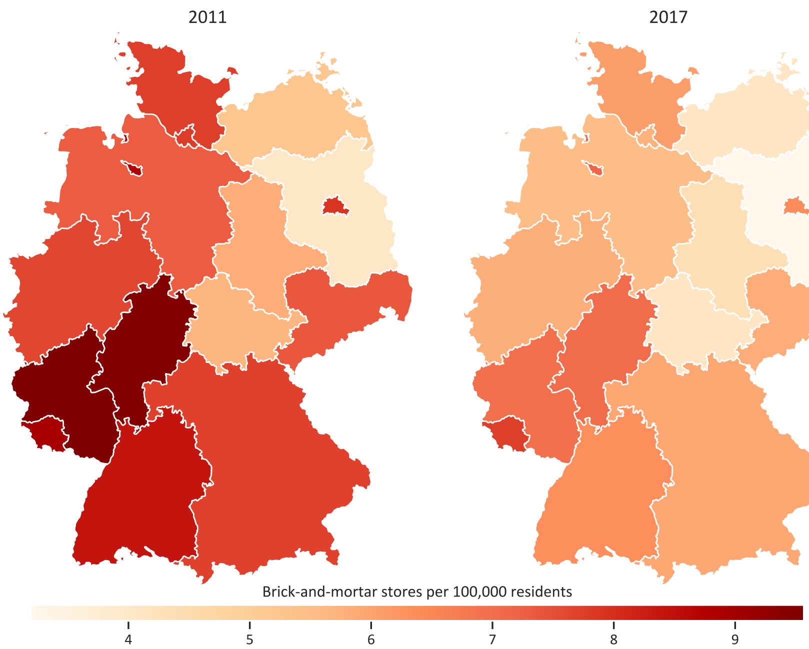


Figure 3: Number of physical bookstores per 100,000 residents in beginning of 2011 and end of 2017. Source: German Publishers and Booksellers Association.

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.3 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.[7]

Figure 4 shows sales figures for physical books from the scanner data

[7]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.

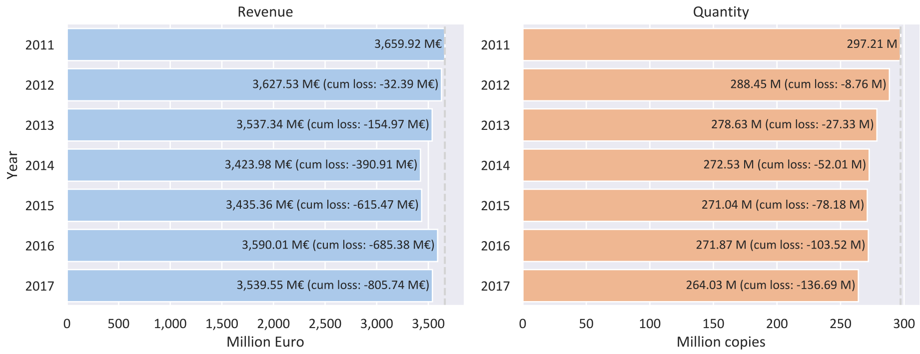


Figure 4: Annual revenues (in million Euros) and sales (in million books sold) of print books in brick-and-mortar stores and e-Commerce. Cumulative losses since 2011 in parentheses. Thuringia and Saarland are excluded due to inconsistencies in the data. Source: media control GmbH.

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 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.3) with the more accurate scanner data, we know

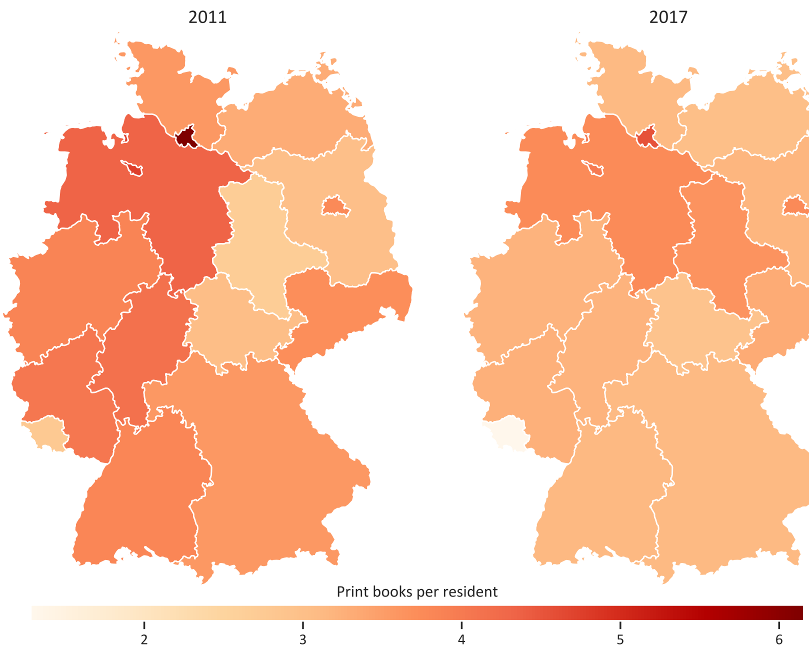


Figure 5: Number of print books sold per resident and year. Source: media control GmbH.

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 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.[8]

In what we presented so far, we have only looked at yearly sales aggregates. To illustrate developments during the years, Figure 7 shows monthly

[8] 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.

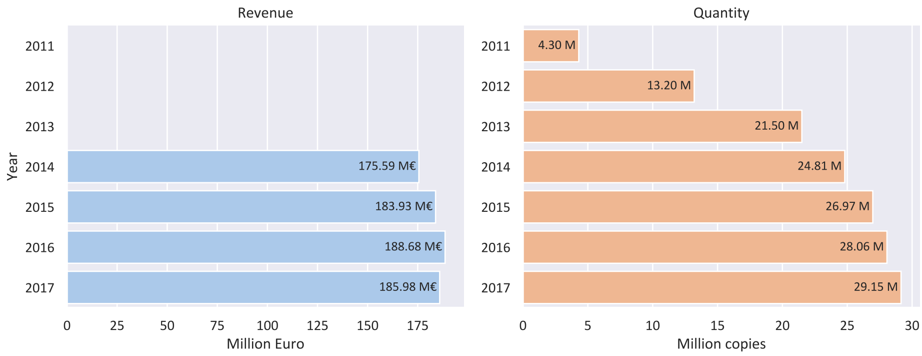


Figure 6: Annual revenues (in million Euros) and sales (in million books sold) of e-Books. Source: GfK GmbH.

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 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 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. During the observation period, there has been a rise in the number of copies sold by individual bookstores. In 2011, the average bookstore sold 3, 630 books per month, resulting in a revenue of 42, 740 Euro. By 2017, this figure increased to 4, 020 books per month, generating 51, 650 Euro in revenue. However, this

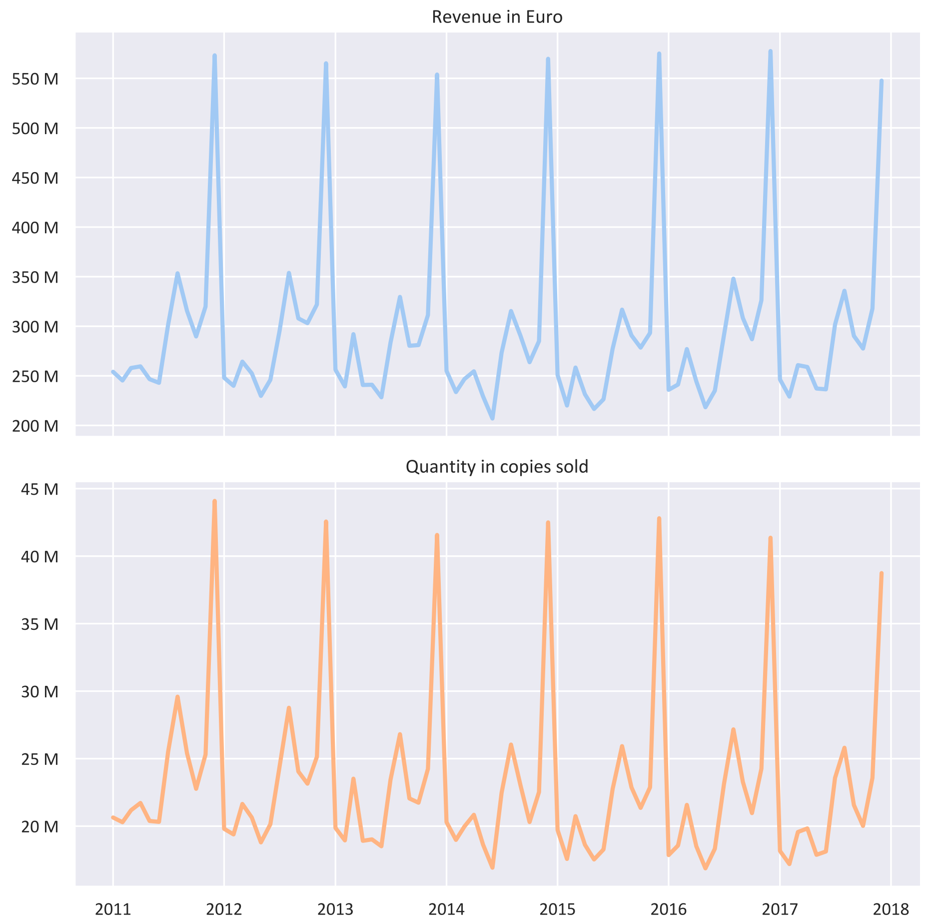


Figure 7: Aggregated sales data (print books) over time (month). Thuringia and Saarland are excluded due to inconsistencies in the data. Source: media control GmbH.

trend is primarily fueled by a rise in market concentration due to some stores exiting the market, rather than an increase in demand.

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

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

percent.[9] 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.

[9]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.

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 (see Appendix A.4 for an description). 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.

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. A prerequisite for (perfect) substitutability between e-Commerce and brick-and-mortar stores is that the drop in the number of physical stores does not have an impact on total book sales because consumers will switch to the e-Commerce channel. A positive relationship between the number of physical stores and book sales instead indicates that e-Commerce and physical stores are imperfect substitutes. With respect to e-Commerce sales, the sign of the coefficients have to be interpreted differently. A positive (negative) coefficient might indicate 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 may also drive the degree to which consumers perceive the offline and the online channel as substitutes. 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 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.[10] 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*

[10] 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.[11]

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

[11]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.[12]

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 ($pop^2_{i,t}$) as an additional instrument.

The relevance condition requires that the instrument is correlated with $\#stores_{i,t}$.[13] 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.[14] 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".[15] Thus, labor force basically captures the share of the population available to

[12]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.

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

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

[15]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.[16]

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

[16]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.[17]

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).[18] We run the regression for total sales, i. e. the sum of sales in physical bookstores and

[17] 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.

[18] 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.[19]

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 different from zero. In that case, depending on which effect is considered, the closure of a bookstore would not affect total sales, 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. It indicates complementarity between the channels, which 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. Even though the exact mechanism cannot be

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

clearly identified based on our data, there are potential explanations for this pattern. One potential explanation could be "showrooming" (Zhang et al., 2018). Although RPM eliminates price differences between the offline and online channel, non-price factors like product fit and wait times, as highlighted by Gensler et al. (2017), can still drive such an effect. Another potential explanation would be word-of-mouth (Beck, 2007), where initial exposure to a title in the offline channel, which is characterized by a much higher sales volume in Germany (see Section A.3 in the Appendix), can lead to increased online purchases. In Section 4.2.2, the effect will be disentangled for different genres and types of bookstores.

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 answer the following question: How did the closures of bookstores affect print book sales in the period 2011-2017? To answer this question, first note that monthly sales went down by 2,765,000 from 24,767,500 in 2011 to 22,002,500 in 2017 (see Figure 4). In the 14 federal states examined in this section, 1,382 bookstores exited the market.[20] With an average decrease in sales upon the closure of a bookstore of 744 units per month on average, store closures have accounted for a drop of

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

1,028,208 in monthly sales. This suggests that bookstore closures accounted for about 37% of the sales decrease.[21]

When interpreting these findings, one has to 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.

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 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

[21] We thank one of the referees for suggesting this interpretation.

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.

Lastly, as a first robustness check[22], we compare the results of our IV-estimation presented above with three alternative specifications: an IV-estimation using yearly data (a), a naive OLS estimation (b) and an OLS estimation with lags (c). The IV-estimation using yearly variation (a) has the benefit that seasonal patterns in sales or the number of bookstores (e.g., the Christmas effect or memberships of the German Booksellers Association running out) play no role in the estimations. However, the number of observations drops from 1,176 to 98, which strongly reduces the power of the analysis. The naive OLS estimation (b) ignores potential endogeneity that arises when book sales are regressed on the number of bookstores (see the beginning of this chapter). The lagged OLS-specification (c) can be seen as a more casual attempt to control for this endogeneity than an IV-approach, since the book sales in period t should not directly affect the number of bookstores in $t + 1$. The results of the respective estimators for $\#stores$ are presented in Table 2.

The second to fourth rows in Table 2 depict the estimation results for the aforementioned, alternative specifications, whereas the first row depicts the results derived from the main IV. In comparison to the latter, the IV with yearly variation (a) shows that the effects of a closure on aggregated sales or sales in e-Commerce are not significantly different from zero at the 5 %-level. This is not surprising given the decrease in the number observations.

[22]Further robustness checks are presented in Section 5.

Table 2: Comparison of IV- and OLS-estimations

	(1) Aggregated	(2) Offline	(3) e-Commerce
IV	0.000248***	0.000183***	0.0000559*
	(0.0000611)	(0.0000486)	(0.0000260)
Yearly	0.00298	0.00589***	0.000565
	(0.00167)	(0.0000521)	(0.000971)
OLS	0.000515	0.000470	0.000106
	(0.000249)	(0.000224)	(0.0000493)
Lagged OLS	0.000557	0.000562*	-0.00000515
	(0.000272)	(0.000222)	(0.0000983)
Federal State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

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

Interestingly, the coefficient of $\#stores$ is still significantly different from zero at the 10 %-level. Moreover, the effects of a closure on offline sales remains significant at the 0.1 %-level. With a naive OLS (b), the coefficient of $\#stores$ is not statistically significant at the 5 %-level. The coefficient is significant at the 10 %-level for all channels (Aggregated, Offline and Online). However, given that the naive OLS does not account for potential endogeneity, the effect may be biased, anyway. The lagged OLS estimates (c) are statistically significant at the 5 %-level for the offline channel and at the 10 %-level in aggregate. The effect in e-Commerce has a p-value of 0.96, so it should not be interpreted.

Overall, it remains to conclude that, at that point, the results of the main IV-estimation appear to be robust. As expected, the significance levels strongly decrease when using yearly data, however, they are in line with what we observe when using monthly data. Also the effects of naive or lagged OLS estimations point in a similar direction as the IV results. It should be noted, though, that the OLS estimations should be interpreted carefully as they are expected to be affected by endogeneity.

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 3 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 3 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 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 finding is in line with expectations as the demand for school books can be expected to be highly inelastic and independent of the number of bookstores. In that genre, service aspects such as expert opinion, ad hoc sales

Table 3: 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.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

etc. should be almost irrelevant.

The coefficients in the third column in Table 3 (Sales e-Commerce) require some discussion. As already explained, these coefficients capture a complementarity between the online and the offline channel in a sense that online sales are positively affected by the number of bookstores. 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).

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.3 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 indicates complementarity between the online and the offline channel.

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. Taking into that between 2011–2017, 1,382 bookstores were closed across the 14 federal states comprising our data set, our results indicate that around 37% of the drop in monthly 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. This finding indicates complementarity between the channels for some genres. On the other hand, we find no such relationship 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.

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.

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.

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.

Gensler, S., Neslin, S. A., & Verhoef, P. C. (2017). The showrooming phenomenon: It's more than just about price. *Journal of Interactive Marketing*, *38*(1), 29–43.

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. (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.

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.

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.

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.

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.

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.

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.

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.

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., & 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.

Williams, R. J. (2024). Empirical effects of resale price maintenance: Evidence from fixed book price policies in europe. *Journal of Competition Law & Economics*, *20*(1-2), 108–136.

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

A.1 Brick-and-mortar bookstores

The numbers presented in Figure 1 are derived from the membership database of the German Publishers and Booksellers Association. Roughly $85 - 90\%$ 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.[A.1] This address book was established in 1839, and is now managed by a subsidiary of the German Publishers and Booksellers Association.[A.2] 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

[A.1] Adressbuch für den deutschsprachigen Buchhandel, see https://adb-online.de; last accessed December 19, 2023.

[A.2] MVB Marketing- und Verlagsservice des Buchhandels GmbH, see https://mvb-online.de; last accessed December 20, 2023.

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. 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.[A.3] 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

[A.3]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).

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 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.

A.2 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.3 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.[A.4]

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.

[A.4] 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.

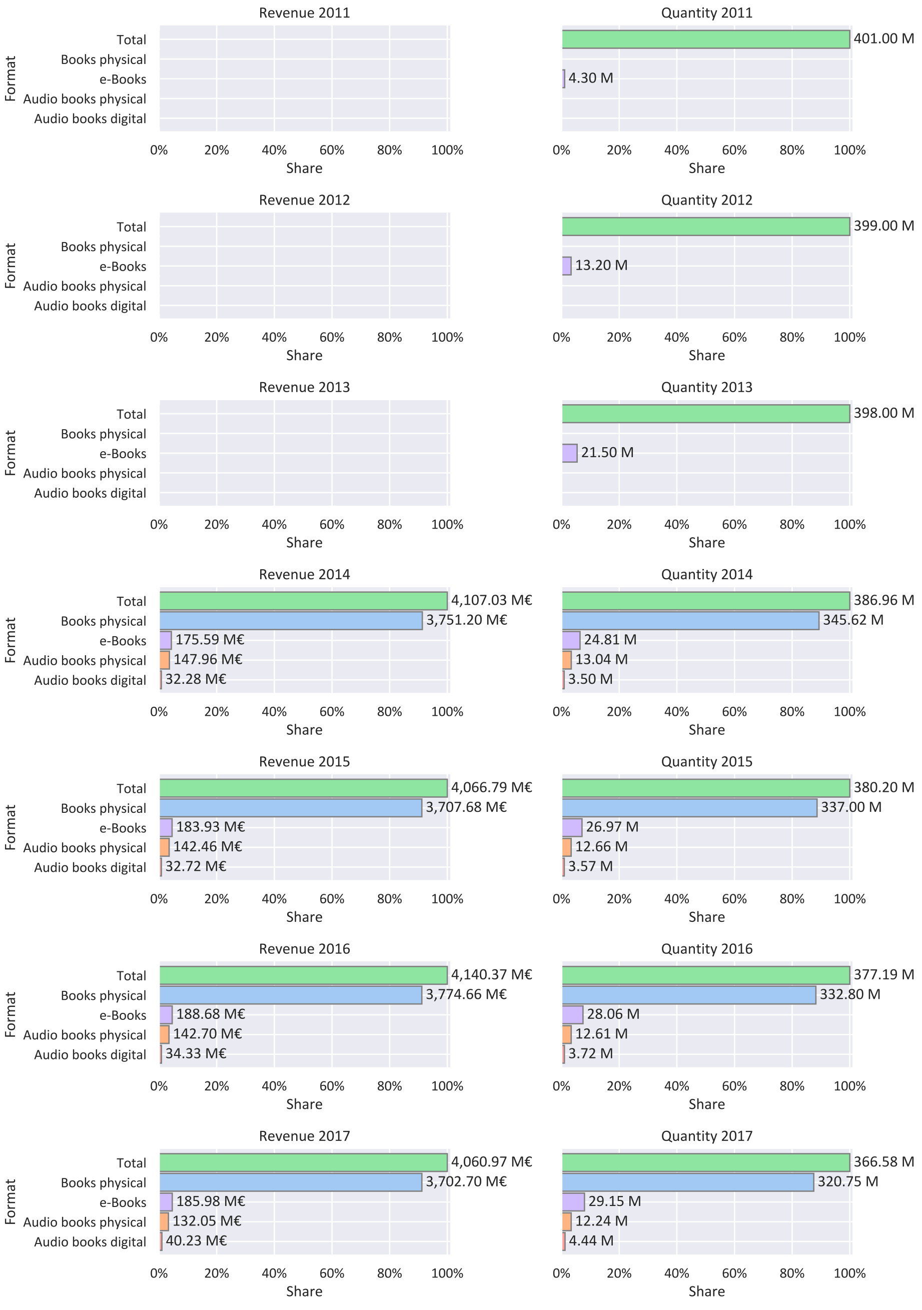


Figure A.1: Evolution of revenue and quantity by format over time (year). Source: GfK GmbH.

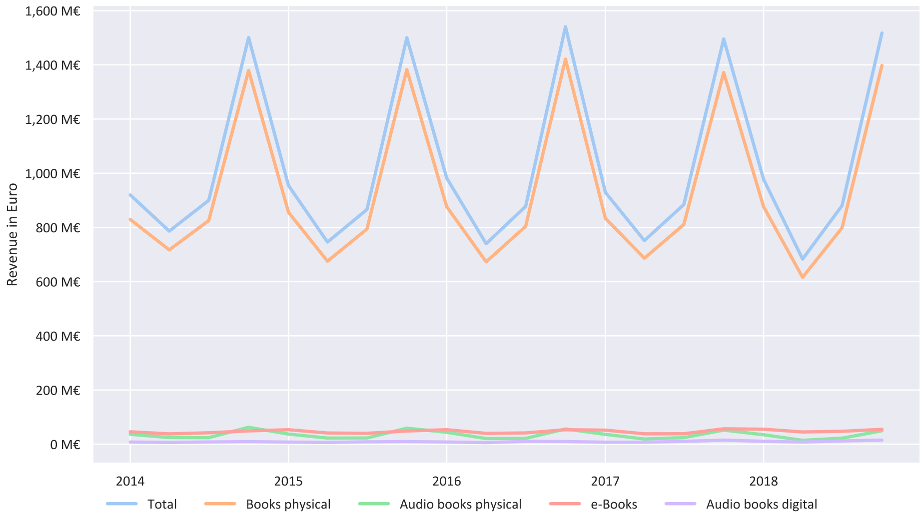


Figure A.2: Evolution of revenues by format over time (quarter). Source: GfK GmbH.

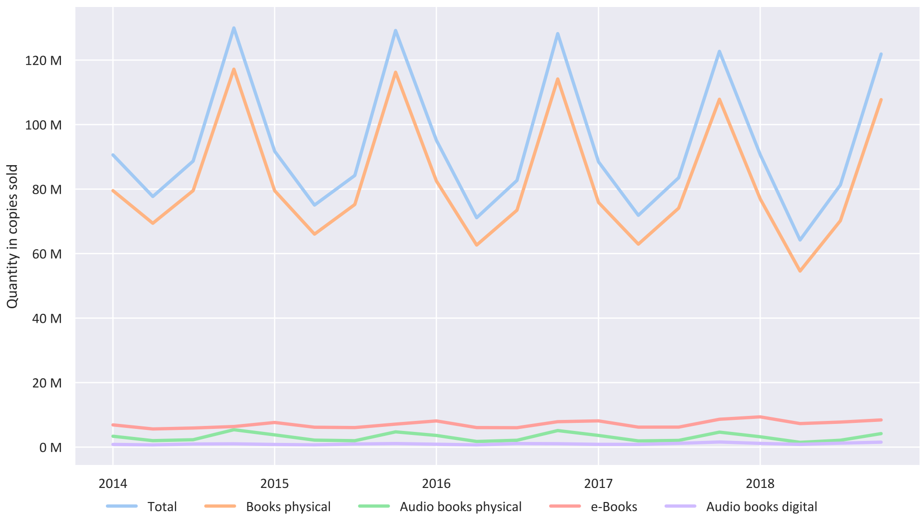


Figure A.3: Evolution of sales volumes by format over time (quarter). Source: GfK GmbH.

A.4 Google Trends 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

51

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.[A.5]

Figure A.4: 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

[A.5] See https://newsinitiative.withgoogle.com/en-gb/resources/trainings/basics-of-google-trends/; last accessed December 26, 2023 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 A.4 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 A.4 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.

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[C.1] 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
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[C.1] 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).

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

C.5 State-specific time trends

Table C.5: State-specific time trends

	(1) Aggregated	(2) Offline	(3) e-Commerce
# Stores	0.000631***	0.000501***	0.000130***
	(0.0000320)	(0.0000261)	(0.00000688)
Federal State FE	Yes	Yes	Yes
Year-Month FE	No	No	No
State-specific time trend	Yes	Yes	Yes
Anderson-Rubin Wald F-statistic	187.5	178.7	171.5
Kleibergen-Paap rk Wald F-statistic	298842.5	298842.5	298842.5
# of observations	1176	1176	1176

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

4. Health Markets

4.1. 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*, vol. 9, no. 7, 2021.

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

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:

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

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.

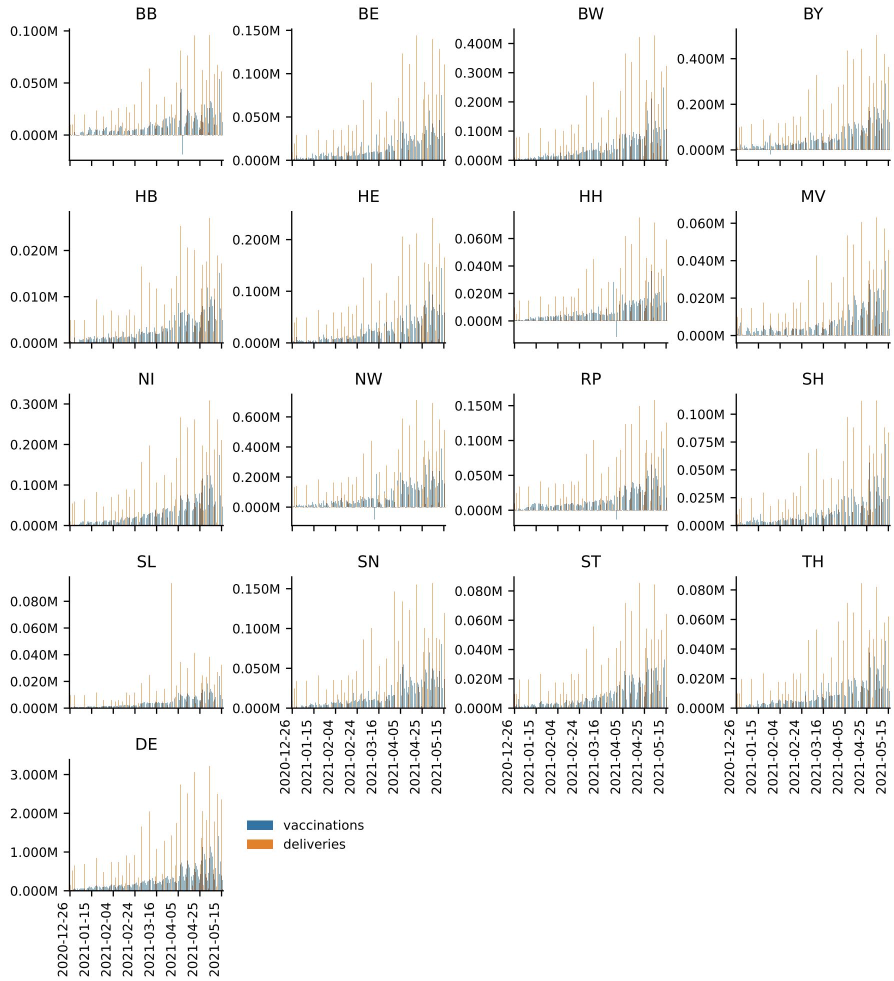


Figure 1. Daily number of people vaccinated and deliveries by Federal State.

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.

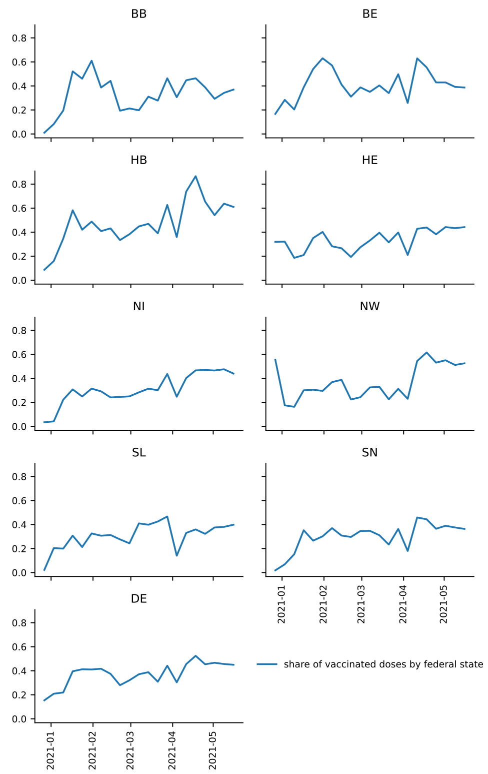


Figure 3. The share of vaccinations given in week t in relation to deliveries in week t and reserves in week $t-1$ $\left(s_{i,t} = \frac{\text{vaccination}_{i,t}}{\text{deliveries}_{i,t}+\text{reserves}_{i,t-1}}\right)$ for all federal states.

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:

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

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.

$$s_{i,t} = \alpha + \beta t + \rho D_{\text{April 5}} + \eta_i D_i \times D_{\text{April 5}} + D_i + \epsilon_{i,t}. \quad \text{(A2)}$$

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:

$$s_{i,t} = \alpha_i + \beta_i t + \rho_i D_{\text{April 5}} + \gamma_i D_{\text{April 5}} \times t. \tag{A3}$$

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]

4.2. Evaluation of a Partial Ban of Rx-Rebates in Germany Using Difference-in-Differences

Authors: Maximilian Maurice Gail, Georg Götz, Daniel Herold and Jan Thomas Schäfer

This paper has been submitted to the journal *Health Policy* and is currently under review.

The manuscript has been presented at the following conference:

- VfS Annual Conference 2022, Verein für Socialpolitik, Basel, Switzerland.

Evaluation of a Partial Ban of Rx-Rebates in Germany Using Difference-in-Differences

Maximilian M. Gail* Georg Götz* Daniel Herold*
Jan Schäfer*

December 2, 2024

Abstract

In December 2020, Germany implemented a policy restricting online pharmacies from offering rebates on prescription drugs to members of the statutory health insurance. This policy change created a natural experiment, allowing us to analyze its impact on the pharmaceutical market using Difference-in-Differences. Utilizing a novel dataset, we find that the ban led to a shift in consumer behavior, increasing offline pharmacy Rx sales by 1.36 % to 1.65 %. However, the policy's effects were unevenly distributed across pharmacies. While all pharmacies experienced some benefit, the impact was disproportionately larger for higher-revenue pharmacies. For instance, pharmacies in the lowest revenue decile saw a modest annual profit increase of € 1,360, whereas those in the highest decile gained more than five times that amount. Our findings indicate that the introduction of VOASG alone was insufficient to reverse the declining trend in pharmacy numbers in Germany. To strengthen the comprehensive supply of pharmaceuticals to the general population, additional reforms are necessary.

Keywords: Pharmacies, Prescription Drugs, Resale Price Maintenance, Regulation, Public Health

JEL Codes: L5, I18

*Chair for Industrial Organization, Regulation and Antitrust, Department of Economics, Justus Liebig University Giessen. Licher Strasse 62, 35394 Giessen. Corresponding author: daniel.herold@wirtschaft.uni-giessen.de.

1 Background

The optimal allocation of healthcare resources, including pharmaceuticals, is a critical public health concern in many countries (Mays et al., 2009; Herwartz and Schley, 2018; Haschka et al., 2020; Li and Liu, 2021). While the specific roles of pharmacies vary across countries, the OECD's characterization of pharmacists as "managing the distribution of medicines to consumers/patients and supporting their safe and efficacious use" provides a generally applicable framework (OECD, 2023). German law (specifically §1 of the Pharmacy Act (ApoG)) requires brick-and-mortar pharmacies to provide the general population with access to medications. However, the number of pharmacies has significantly decreased, dropping by roughly 12.5% from 2010 (21,441) to 2020 (18,753) (ABDA, 2024, p. 9). This decline coincides with a rise in competition from foreign online pharmacies. Their market share for over-the-counter (OTC) drugs increased from around 5% in 2008 to 20% by 2020 (ABDA, 2021; Statista, 2024). Initially, resale price maintenance (RPM) for prescription (Rx) drugs limited price competition to OTC medications. However, a 2016 European Court of Justice ruling allowed online pharmacies to also offer discounts on Rx drugs (Albrecht et al., 2020).

In the light of these developments, in December 2020 the German government implemented a law that was supposed to strengthen the comprehensive supply of pharmaceuticals to the general population by supporting brick-and-mortar pharmacies. Accordingly, the name of the law was "Vor-Ort-Apotheken Stärkungsgesetz" (Local Pharmacy Support Act), henceforth referred to as VOASG. This law partially restricts online rebates by prohibiting them for Rx drugs sold to members of the statutory health insurance scheme, but allows them for privately insured individuals (details on the German health insurance system in Section 2.2). This article investigates the impact of the VOASG on sales and rents in the Rx market using a difference-in-differences (DiD) approach. By comparing the dispensation of Rx drugs to members of the statutory health insurance (treatment group) with those to the privately insured and self-pay patients (control group), we can isolate the effect of the policy change. The VOASG only affected Rx drugs dispensed to statutory health insurance members, leaving prescriptions to the privately insured and self-pay patients unaffected. This differential impact allows us to identify the causal effect of the policy.

We employ a novel dataset for this study, constructed from high-frequency sales data provided by the major merchandise information system (MIS) suppliers in Germany. This dataset encompasses individual transaction data from approximately 9,231 offline pharmacies, representing nearly half of all German pharmacies, for the period January 1, 2018, to October 31, 2022 (see Section 2.1).

We find that the partial ban on rebates led to an increase in offline sales of around 1.36 to 1.65 % for an average brick-and-mortar pharmacy compared to a counterfactual scenario in which rebates would not have been banned. Given that the demand for Rx drugs can be considered price inelastic, any increase in offline sales should correspond to a decrease in online sales by the same amount.

Our findings reveal that a substantial portion of consumers respond to price differences and rebates for Rx drugs by selecting the more affordable retail channel. This insight is particularly relevant in light of the 2016 European Court of Justice ruling (Case No.: C-148/15) mentioned above, which legalized rebates offered by foreign online pharmacies. In this ruling, the judges claimed that the German government had failed to show that RPM was an effective tool to achieve the alleged goal of securing the comprehensive supply of pharmaceuticals to the general population. Our research provides empirical evidence that consumers are price-sensitive, suggesting that online rebates could potentially erode offline sales and profitability.

Given these findings, we further investigate whether the policy change successfully mitigated large-scale pharmacy closures. Economic theory posits that market exit occurs when opportunity costs exceed revenues, resulting in negative economic profits (Jehle and Reny, 2011, Ch. 4). Ceteris paribus, pharmacies with lower revenues are more susceptible to market exit. To evaluate the impact of VOASG, we stratified the sample into revenue deciles and calculated the DiD effect for each decile.

Our analysis reveals that the effects for the lowest seven deciles are relatively similar, while the effect in the three highest deciles is around 40% to 140% stronger. This suggests that larger pharmacies benefited disproportionately from the rebate ban compared to smaller pharmacies. These findings indicate that the law, against its stated goal, did not significantly support pharmacies at risk of market exit. We estimated the additional *annual* profits generated by the rebate ban for pharmacies in the first and tenth deciles, which were approximately €1,360 and €7,690, respectively. Notably, the *additional annual profit* for the lowest decile equates to around one third of the average *monthly* income of an employee in Germany in 2021 (Destatis, `https://t.ly/N2dJq`). The findings also indicate that the majority of pharmacies experienced only a mild increase in profits, with a median increase in profits of €3,246. Given these relatively modest gains, it seems unlikely that the policy had a substantial impact on the market. This is corroborated by the developments in 2021-2023, when another 6.3% of pharmacies closed.

This article contributes, first, to the growing body of research on public health service provision. While many studies have focused on hospitals or physicians in specific countries such as the US (Mays et al., 2009; Duminy et al., 2022), China (Li and Liu, 2021), and Germany (Herwartz and Schley, 2018; Haschka et al., 2020), our research examines the role of pharmacies. These establishments often serve as the final link in the pharmaceutical supply chain, delivering medications to the general public (Inoue et al., 2016; Raza et al., 2022). Previous research on pharmacies has primarily investigated specific services, such as their role in delivering primary care or providing non-prescription medications (Smith, 2009; Agomo, 2012; Perraudin et al., 2016; Costa et al., 2019). Our research, in contrast, analyzes how price competition affects pharmacy profitability and, consequently, the financial sustainability of their services. By examining this aspect, our study adds a new dimension to the understanding of pharmacy's role in public health.

Second, this article contributes to the literature on digitization, particularly the debate surrounding the substitutability of offline and online services (Brynjolfsson and Smith, 2000; Brown and Goolsbee, 2002; Sinai and Waldfogel, 2004; Jin and Kato, 2007; Goldmanis et al., 2010; Cavallo, 2017; Couture et al., 2021). While research on digital public health explores how digitization can improve population health (Iyamu et al., 2022; Wong et al., 2022; Yurkovich et al., 2024), our study focuses on the pharmacy sector. Specifically, we investigate the extent to which online pharmacies complement or substitute traditional brick-and-mortar pharmacies in supplying pharmaceuticals to the public (Coenen et al., 2011; an der Heiden and Meyrahn, 2017). Our findings suggest that price differentials impact consumer behavior. When price disparities exist between online and offline channels, there is a statistically and economically significant fraction of consumers that chooses the cheaper option. This price sensitivity has important implications for policymakers aiming to maintain a network of brick-and-mortar pharmacies of a certain density. As online pharmacies, especially those offering rebates, can cannibalize the market share of traditional pharmacies, policymakers should consider the potential economic consequences of price competition.

Third, we contribute to the research on RPM (Telser, 1960; Marvel and McCafferty, 1985; Hunold and Muthers, 2017). While RPM can potentially eliminate freeriding by online retailers and mitigate double marginalization, it also suppresses downstream price competition (Elzinga and Mills, 2008). In the pharmacy market, the situation is more complex. Countries often pursue health policy objectives that extend beyond market-based outcomes, such as ensuring a certain number or quality of pharmacies (Wambach et al., 2018). Our findings indicate that RPM shifts profits towards brick-and-mortar pharmacies, thereby contributing to the achievement of these objectives. However, the introduction of VOASG alone was insufficient to reverse the declining trend in pharmacy numbers in Germany. Other structural factors (in particular, pharmacies' remuneration) seem to exert a more substantial influence on the decline of offline pharmacies than RPM.

The article is structured as follows. In Section 2, we discuss the data and our identification strategy. Section 3 contains the results of our empirical analysis. In Section 4, we discuss how the identified effects differ between pharmacies with different revenues as well as limitations of the study. Section 5 concludes.

2 Methods

In this section, we outline the methodological approach employed to identify the impact of the partial ban on online Rx rebates on the market. We begin in Section 2.1 by describing the dataset we assembled for our analysis. Subsequently, in Section 2.2, we provide a non-technical overview of our identification strategy. A more technical implementation of our identification strategy is detailed in Appendix A.

2.1 Data

In our empirical analysis, we use sales data for 5,487 pharmacies. This sample is derived from a larger dataset of 9,231 pharmacies, which we adjusted for our analysis as described in Appendix A.2. The final sample represents 29.13% of the total number of 18,839 pharmacies as of September 29, 2020 (see Figure B.1 for a map of the geographical distribution).

Through the balancing procedure and the application of these data constraints, we retain approximately 68.2% of the total sales data from the unrestricted dataset. Ultimately, we assemble two distinct annual datasets, ranging from 2018 to 2021: one categorized by 2-digit zip code and another by individual pharmacy. Each dataset contains comprehensive transaction details for each pharmacy, including sales volumes, the AVP (retail price), and other product-specific information. For further details on the data processing steps, including balancing, constraints, and aggregation, refer to Appendix A.2.

2.2 Identification Strategy

In this section, we outline our identification strategy, which leverages the differential impact of the VOASG on specific population segments. This natural experiment allows us to apply a DiD approach and conduct an event study to assess causal effects (Cunningham, 2021, Chapter 9). Prescriptions to members of the statutory health insurance, directly affected by the VOASG, form the treatment group, while prescriptions to privately insured individuals and self-pay patients, unaffected by the reform, serve as the control group. By comparing the differential changes in the dispensation of Rx drugs between these groups, we isolate the causal effect of the VOASG. To provide further clarity, we first offer a brief overview of the German insurance system, with a particular focus on the prescription drug dispensation scheme.

In Germany, prescription drugs are prescribed by physicians and dispensed by both traditional and online pharmacies. To access these medications, members of the statutory health insurance are typically required to make co-payments based on AVP, contributing a portion of the drug's cost. Private insurance offers a range of reimbursement schemes, often involving initial out-of-pocket expenses and subsequent reimbursement by the insurer. (For a more detailed explanation, see Section C.2.) Thus, both systems generally adhere to standardized reimbursement rates for prescribed medications. Irrespective of the specificities of each insurance scheme, drug prices are the same in the offline and online channels irrespective of insurance. The same is true for self-pay patients. The only difference in the price dimension is that, after the introduction of VOASG, members of the statutory health insurance were no longer entitled to rebates for online purchases.

The VOASG came into effect on December 15, 2020. As described above, a central aspect of this legislation is the prohibition of foreign online pharmacies from offering rebates to individuals insured under the statutory health insurance. However, these pharmacies are still permitted to provide rebates to privately insured consumers (Federal Ministry of Health, `https://www.bundesgesundheitsministerium.de/apotheken.html`).

Rebates granted by online pharmacies usually range from €2.50 to €10 and take the form of vouchers. In relation to the fundamental differences between public and private health insurance (premium structures, reimbursement mechanisms, service quality) these rebates are of minor importance.

There are also notable barriers to switching between the two systems. Individuals usually can only switch when their employment status changes in a special way in terms of income and type of employment. The policy can thus be considered exogenous to the individuals' choice regarding insurance schemes.

The demand for prescription drugs is generally inelastic due to their nature (Gatwood et al., 2014; Yeung et al., 2018): a patient's need for medication is often diagnosed by a physician and is not easily deferred. Moreover, as the patient's insurance typically covers the majority of the cost, co-payments are of minor importance. This suggests that patients are unlikely to forego necessary medication solely due to the absence of rebates. While rebates potentially influence a patient's choice of pharmacy (online or offline), their impact on overall drug consumption is likely minimal. Given the inelastic nature of demand, any shift in offline sales is expected to be accompanied by a similar shift in online sales.

Given that pharmacy compensation is directly tied to the number of packages dispensed, we utilize this metric to assess the impact of VOASG. This approach is further justified by the fact that market data is predominantly reported in terms of sales or revenue, eliminating the need for additional conversions. A more in-depth discussion is presented in the following section.

3 Results

To gain an initial understanding of a potential DiD effect, we compared the mean sales of the treatment and control groups before and after the introduction of VOASG. Post-VOASG, prescription drug sales in the treatment group exhibited a 1.63% increase relative to the control group. A detailed summary of this finding is provided in Appendix B.1 (Figure B.2), together with further descriptive statistics.

To estimate the causal impact of VOASG, we use a two-way fixed effects (TWFE) DiD estimation. Additionally, an event study is conducted to assess the common trends assumption by examining pre-treatment periods before the implementation of VOASG (Cunningham, 2021, Chapter 9.4). Technical details on the estimation procedures can be found in Appendix B.

As covariates, we include (i) the weighted quantity of doses of Rx drugs dispensed and (ii) the fraction of customers that are members of the statutory health insurance.

Covariate (i) is calculated based on Germany's N-classification system. This system categorizes pharmaceutical packages into three sizes: N1 (10 doses), N2 (30 doses), and N3 (100 doses). While not exact, this system provides a reasonable approximation of package sizes. To account for potential variations in package sizes that could influence our results, we calculate a weighted average of package sizes at the 2-digit zip code or pharmacy level for each group. For instance, a combination of 15 N1 and 5 N2 packages would equate to approximately 300 doses, which corresponds to a weighted average of 20 doses. By controlling for this covariate, we mitigate the impact of package size differences on our analysis (see Appendix A.2 for more details).

Covariate (ii) accounts for the relative share of individuals enrolled in statutory and private health insurance. This covariate is necessary to mitigate the potential confounding effects of significant shifts in insurance enrollment. Due to data constraints, this covariate is measured at the national level and varies annually, essentially functioning as a time trend.

Table 1 presents the results of the TWFE DiD model (see Equation (B.1) in Appendix B). We estimate two model specifications: a baseline model (A) and a model with a time-trend (B). For each specification, Column (1) presents results aggregated at the 2-digit zip code level with robust standard errors. Columns (2) to (4) reports the results for an aggregation at pharmacy-level. Column (2) uses robust standard errors, while Columns (3) and (4) cluster standard errors at the 2-digit zip code level, or at the 2-digit zip code and insurance group level, respectively.

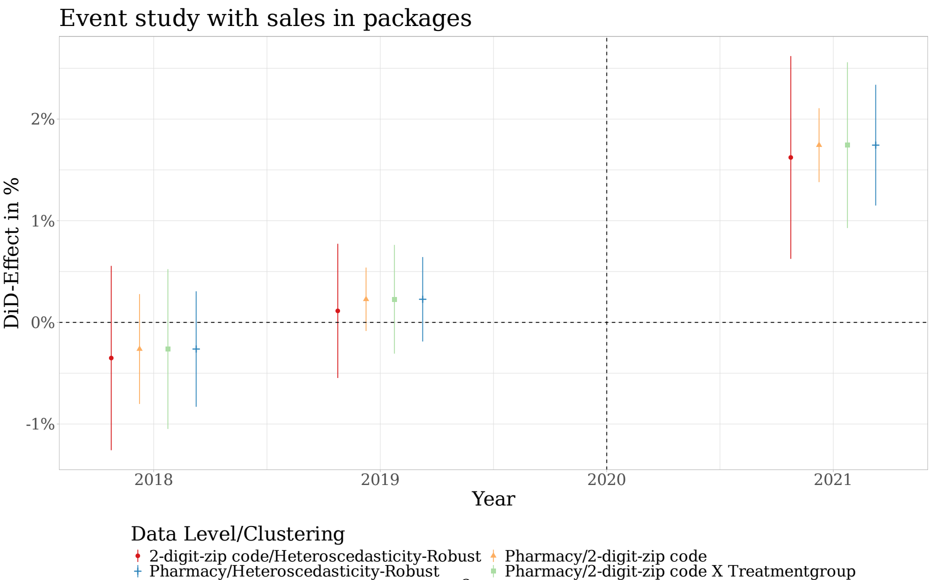
Table 1 shows a DiD effect that ranges from 0.0136 to 0.0165. All results are statistically significant at the 1 % level. These effects represent the average treatment effect on the treated (ATT) and can be interpreted as percentage changes. Therefore, due to the introduction of the VOASG, sales increased by approximately 1.36 % to 1.65 % compared to a counterfactual scenario without the policy. Further discussion of the result follows in Section 4.

The remainder of this section presents the results of an event study approach. The technical details are presented in Appendix B.

	(1)	(2)	(3)	(4)
Specification A: Sales in Packages				
DiD-Coefficient	0.0155***	0.0165***	0.0165***	0.0165***
	(0.0039)	(0.0023)	(0.0015)	(0.0032)
Fraction of Members in Insurance	1.5718	1.1123	1.1123	1.1123
	(1.8638)	(1.2091)	(1.1569)	(1.7127)
Weighted Average of Doses	0.0013	-0.0006	-0.0006	-0.0006
	(0.0030)	(0.0006)	(0.0009)	(0.0008)
Observations	760	43,896	43,896	43,896
Adj. R2	0.9997	0.9938	0.9938	0.9938
FE: Year	X	X	X	X
FE: 2 digit zip code & Treated	X			
FE: Pharmacy & Treated		X	X	X
Std. Errors	Robust	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated
Specification B: Sales in Packages with Trends				
DiD-Coefficient	0.0136***	0.0149***	0.0149***	0.0149***
	(0.0048)	(0.0029)	(0.0018)	(0.0036)
Weighted Average of Doses	0.0012	-0.0006	-0.0006	-0.0006
	(0.0030)	(0.0006)	(0.0009)	(0.0008)
Time Trend of Treatment-Group	0.0017	0.0013	0.0013	0.0013
	(0.0017)	(0.0011)	(0.0010)	(0.0015)
Observations	760	43,896	43,896	43,896
Adj. R2	0.9997	0.9938	0.9938	0.9938
FE: Year	X	X	X	X
FE: 2 digit zip code & Treated	X			
FE: Pharmacy & Treated		X	X	X
Std. Errors	Robust	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated

Note:
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 1: Results of the DiD-estimation.



9

Figure 1: Event study results with 99 % confidence intervals.

The event study provided in Figure 1 highlights that all pre-treatment coefficients are statistically insignificant and close to zero. The ATT for 2021 is estimated at 0.0162 and 0.0174, maintaining statistical significance similar to the DiD estimation. Thus, results from the event study align closely with those of the regression analyses above (see Figure B.2).

As explained in Section 2.2, we measure drug dispensation in terms of packages. To account for variations in package sizes, we employ a covariate based on the N-classification system, as described above. As a robustness check, we conduct additional analyses using approximated dosages derived from the N-classification system. Further details and results are provided in Appendix A.2.

4 Discussion

In this section, we discuss our findings in the context of the policy objective, which aimed to support (small) brick-and-mortar pharmacies. The section is divided into two parts. In Section 4.1, we extend our analysis to examine how pharmacies with different revenue levels benefited from the partial ban on rebates. In Section 4.2, we address the limitations inherent to our study.

4.1 Distributional Effects

As explained in Section 1, the VOASG was introduced to strengthen the comprehensive supply of pharmaceuticals to the general population during a period of a drastically shrinking number of brick-and-mortar pharmacies. To evaluate the effect of the policy change on the number of pharmacies in the market, we extend the analysis presented in Section 3.

As explained in the introduction, all else equal, pharmacies with lower revenues are more exposed to market exit. We therefore categorize pharmacies into deciles based on their revenue. The first decile includes the 10% of pharmacies with the lowest revenue, the second decile includes the next 10% with the second-lowest revenue, and so on. We then interact the DiD effect estimated in Section 3 with a dummy variable indicating each decile. A time trend is included to control for potentially diverging evolutions between the two insurance groups. The results of this analysis are presented in Table 2. Technical details are outlined in the Appendix (Section C.3, Equation (C.1)).

	(1)	(2)	(3)
Weighted Average of Doses	-0.0017***	-0.0017**	-0.0017**
	(0.0005)	(0.0007)	(0.0007)
Time Trend of Treatment-Group	0.0016*	0.0016	0.0016
	(0.0009)	(0.0010)	(0.0013)
DiD for Decile 1	0.0128***	0.0128**	0.0128**
	(0.0047)	(0.0053)	(0.0056)
DiD for Decile 2	0.0111***	0.0111***	0.0111**
	(0.0043)	(0.0038)	(0.0046)
DiD for Decile 3	0.0107***	0.0107**	0.0107**
	(0.0040)	(0.0041)	(0.0043)
DiD for Decile 4	0.0146***	0.0146***	0.0146***
	(0.0042)	(0.0041)	(0.0044)
DiD for Decile 5	0.0127***	0.0127***	0.0127***
	(0.0040)	(0.0034)	(0.0044)
DiD for Decile 6	0.0092**	0.0092**	0.0092**
	(0.0037)	(0.0036)	(0.0040)
DiD for Decile 7	0.0131***	0.0131***	0.0131***
	(0.0037)	(0.0030)	(0.0036)
DiD for Decile 8	0.0182***	0.0182***	0.0182***
	(0.0038)	(0.0030)	(0.0039)
DiD for Decile 9	0.0208***	0.0208***	0.0208***
	(0.0037)	(0.0031)	(0.0037)
DiD for Decile 10	0.0190***	0.0190***	0.0190***
	(0.0040)	(0.0039)	(0.0047)
Observations	43,896	43,896	43,896
Adj. R2	0.9959	0.9959	0.9959
FE: Year	X	X	X
FE: Pharmacy & Treated	X	X	X
FE: Deciles	X	X	X
Std. Errors	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated

Note:
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Results of the DiD-estimation for revenue deciles.

The panels presented in Table 2 are defined in Section 3. The coefficients for "DiD for Decile x" capture the interaction between the DiD effect and the respective decile. These coefficients should be interpreted in the same manner as the DiD coefficient in Table 1, but are specific to each decile. For instance, a coefficient value of 0.0128 for "DiD for Decile 1" indicates that, on average, sales of brick-and-mortar pharmacies in the lowest revenue decile (treatment group) increased by 1.28 %, holding all else constant.

The results reveal a notable asymmetry in the magnitude of the effects across different revenue deciles. Specifically, the increase in Rx sales for deciles one to seven ranges from approximately 0.92 % to 1.46 % across all panels in Table 2. In contrast, the effect for the highest three deciles is 40 % to 140 % larger, ranging from 1.82 % to 2.16 %. In other words, pharmacies with higher revenues seem to have benefited more from the partial ban on rebates than those with lower revenues. This disparity is further accentuated by the fact that the results in Table 2 are expressed in relative terms. Given that higher-revenue pharmacies typically have larger absolute sales, the impact in absolute terms is even more pronounced. This can be visualized in Figure 2, which illustrates the absolute increase in remuneration for each decile, calculated based on the results in Table 2.

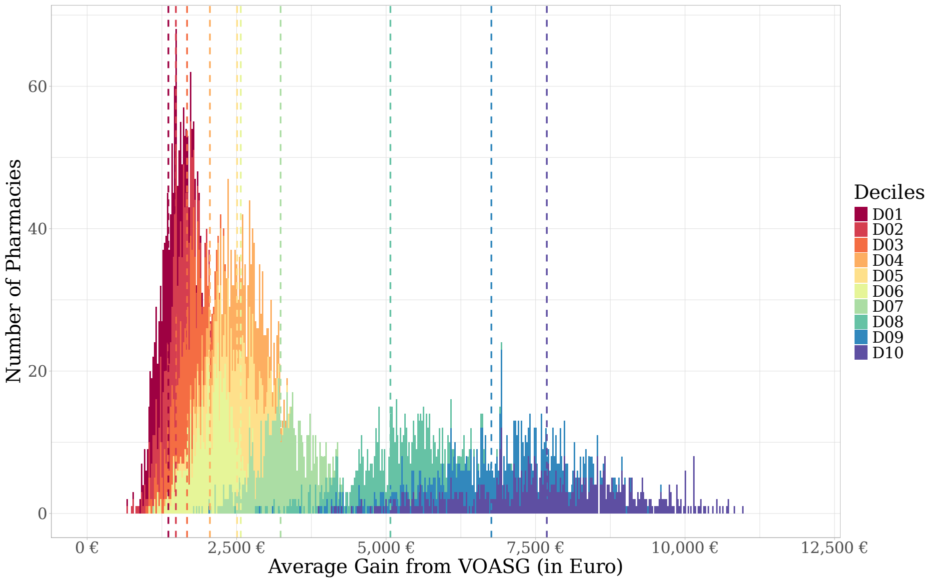


Figure 2: DiD effect for each decile (see Table 2). $Average\ Gain\ from\ VOASG = Sales \times \frac{\beta_{Decile}}{1+\beta_{Decile}} \times Average\ Remuneration\ Per\ Sale.$

Figure 2 illustrates the distribution of the absolute impact of the partial ban on rebates across pharmacies with varying revenue levels. This figure was constructed by multiplying each pharmacy's remuneration for dispensing Rx drugs by its corresponding DiD effect (see Table 2). For instance, the Rx remuneration of the pharmacy in the lowest revenue decile is multiplied by the "DiD for Decile 1", while the remuneration of the highest-revenue pharmacy is multiplied by the "DiD for Decile 10". (For a more detailed explanation of how remuneration is computed, see Appendix C.2.)

The dotted lines in Figure 2 represent the average increase in profits for each decile. As previously discussed, the first seven deciles exhibit significantly lower gains compared to the highest three. Our findings suggest that, on average, the rebate ban generated additional annual profits for pharmacies in the first decile of around €1,360. In contrast, the corresponding gain for pharmacies in the tenth decile is more than five times greater, reaching €7,690. Figure 2 also indicates that 50 % of the pharmacies experienced an increase in profits of less than €3,246 (median). In contrast, the average gain across all pharmacies is at 3,500. These results show that the gains from the introduction of VOASG were unevenly distributed across pharmacies, with *larger* pharmacies actually benefiting more strongly.

These results can be further contextualized. As discussed previously, pharmacies exit the market when opportunity costs exceed revenues. The question, therefore, is whether the additional profits are meaningful enough to sustain pharmacies in the market. To evaluate this, consider that the average *monthly* income of a German employee in 2021 was €4,100 (Destatis, `https://t.ly/N2dJq`). In contrast, the *annual* gain for pharmacies in the first decile approximates one-third of this amount. The majority of pharmacies experienced an increase equivalent to 50-60% of this figure. In other words, pharmacy owners' yearly gains were substantially less than the average monthly wage of an employee. Given that these owners are highly skilled professionals, the impact of the VOASG on their revenues can be considered relatively low.

Against the backdrop of these findings, the continued decline in the number of brick-and-mortar pharmacies after 2020 is unsurprising. By the end of 2023, the total number had decreased to 17,571, reflecting a 6.3 % reduction compared to 2020. While external factors, such as the war in Ukraine and the pandemic, influenced the market, the policy change appears to have failed to address the underlying causes of pharmacy closures.

4.2 Limitations

Our study is subject to certain limitations. While it relies on the most comprehensive and detailed dataset available on the German pharmacy market (to our knowledge), the data is limited to the 2-digit zip code level. Consequently, we are unable to further specify the precise location of a given pharmacy. If the data were more granular, we could have conditioned the effects on specific socio-geographic factors, such as rural versus urban areas and the specific competitive landscape of offline pharmacies.

A comparable dataset encompassing online sales is currently unavailable. Our interpretation, which posits that e-commerce experiences losses equivalent to the gains of the offline channel, is thus contingent on the assumption of inelastic demand.

It is essential to consider these limitations when interpreting our findings. However, these caveats do not affect the validity of our identification strategy.

5 Conclusion

Our study investigated the impact of a partial ban on rebates for prescription drugs (VOASG) on the German pharmacy market. We found that the ban led to a modest increase in offline sales (1.36-1.65 %) for brick-and-mortar pharmacies, likely at the expense of online sales. This suggests that a notable portion of consumers is price-sensitive for Rx drugs and will switch channels for cheaper options.

However, the policy appears to have had an asymmetric effect on pharmacies. We estimated that annual additional profits ranged from €1,360 - €7,690, with a median increase of €3,246. Pharmacies in the top three revenue deciles experienced a 40-140 % stronger sales increase compared to smaller ones.

This observation, coupled with the continued market exit observed after 2021 (6.3 % closure rate), suggests that further reforms are necessary to reverse the trend of declining pharmacy numbers. Such reforms can, for instance, specifically target pharmacy remuneration for dispensing Rx drugs.

References

ABDA, 2021. Zahlen, daten, fakten 2021. https://www.pharma4u.de/fileadmin/user_upload/pdf/ABDA-Leitlinien/ABDA_ZDF_2021_Broschuere.pdf.

ABDA, 2024. Zahlen, daten, fakten 2024. https://www.abda.de/fileadmin/user_upload/assets/ZDF/Zahlen-Daten-Fakten-24/ABDA_ZDF_2024_Broschuere.pdf.

Agomo, C.O., 2012. The role of community pharmacists in public health: a scoping review of the literature. Journal of Pharmaceutical Health Services Research 3, 25–33.

Albrecht, M., Baake, P., an der Heiden, I., Brenck, A., Ochmann, R., Schiffhorst, G., 2020. Ökonomisches Gutachten zum Apothekenmarkt. Technical Report. IGES Institut und Deutsches Institut.

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

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

Cavallo, A., 2017. Are online and offline prices similar? evidence from large multi-channel retailers. American Economic Review 107, 283–303.

Coenen, M., Haucap, J., Herr, A., Kuchinke, B., 2011. Wettbewerbspotenziale im deutschen Apothekenmarkt. Technical Report. DICE - Ordnungspolitische Perspektiven Nr. 17.

Costa, S., Cary, M., Helling, D.K., Pereira, J., Mateus, C., 2019. An overview of systematic reviews of economic evaluations of pharmacy-based public health interventions: addressing methodological challenges. Systematic reviews 8, 1–20.

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

Cunningham, S., 2021. Causal inference: The mixtape. Yale university press.

Duminy, L., Ress, V., Wild, E.M., 2022. Complex community health and social care interventions–which features lead to reductions in hospitalizations for ambulatory care sensitive conditions? a systematic literature review. Health Policy 126, 1206–1225.

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 .

Freyaldenhoven, S., Hansen, C., Shapiro, J.M., 2019. Pre-event trends in the panel event-study design. American Economic Review 109, 3307–38.

Gatwood, J., Gibson, T.B., Chernew, M.E., Farr, A.M., Vogtmann, E., Fendrick, A.M., 2014. Price elasticity and medication use: cost sharing across multiple clinical conditions. Journal of Managed Care Pharmacy 20, 1102–1107.

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

Haschka, R.E., Schley, K., Herwartz, H., 2020. Provision of health care services and regional diversity in germany: Insights from a bayesian health frontier analysis with spatial dependencies. The European Journal of Health Economics 21, 55–71.

an der Heiden, I., Meyrahn, F., 2017. Ermittlung der Erforderlichkeit und des Ausmaßes der in der Arzneimittelpreisverordnung (AMPreisV) geregelten Preise. Technical Report. 2HM & Associates GmbH. URL: `https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/ermittlung-der-erforderlichkeit-und-des-ausmasses-von-aenderungen-der-in-der-arzneimittelpreisverordnung.pdf?__blob=publicationFile&v=12`.

Herwartz, H., Schley, K., 2018. Improving health care service provision by adapting to regional diversity: an efficiency analysis for the case of germany. Health Policy 122, 293–300.

Hunold, M., Muthers, J., 2017. Resale price maintenance and manufacturer competition for retail services. The RAND Journal of Economics 48, 3–23.

Inoue, Y., Takikawa, M., Morita, Y., Takao, K., Kanamoto, I., Sugibayashi, K., 2016. A comparison of pharmacists' role functions across various nations: the importance of screening. Research in Social and Administrative Pharmacy 12, 347–354.

Iyamu, I., Gómez-Ramírez, O., Xu, A.X., Chang, H.J., Watt, S., Mckee, G., Gilbert, M., 2022. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digital health 8, 20552076221102255.

Jehle, G., Reny, P., 2011. Advanced microeconomic theory, 3rd edition. Pearson Education Limited.

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

Li, L., Liu, Z., 2021. Research on efficiency measurement and spatiotemporal disparity of rural public health services in china. PLoS One 16, e0252871.

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

Mays, G.P., Smith, S.A., Ingram, R.C., Racster, L.J., Lamberth, C.D., Lovely, E.S., 2009. Public health delivery systems: evidence, uncertainty, and emerging research needs. American Journal of Preventive Medicine 36, 256–265.

OECD, 2023. Health at a Glance. Technical Report. Organisation for Economic Co-operation and Development.

Perraudin, C., Bugnon, O., Pelletier-Fleury, N., 2016. Expanding professional pharmacy services in european community setting: is it cost-effective? a systematic review for health policy considerations. Health policy 120, 1350–1362.

Raza, M.A., Aziz, S., Noreen, M., Anjum, I., Raza, S.M., 2022. A portrait of the pharmacy profession globally: pharmacist universal professional identity and establishment of global pharmacy council. INNOVATIONS in pharmacy 13.

Roth, J., Sant'Anna, P.H., Bilinski, A., Poe, J., 2023. What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Journal of Econometrics 235, 2218–2244.

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

Smith, F., 2009. Private local pharmacies in low-and middle-income countries: a review of interventions to enhance their role in public health. Tropical Medicine & International Health 14, 362–372.

Statista, 2024. Absatzanteile des arzneimittelversandhandels am deutschen otc-markt nach produktkategorie in den jahren 2008 bis 2019. https://de.statista.com/statistik/daten/studie/280477/umfrage/absatzanteile-am-otc-versandapothekenmarkt/.

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

Wambach, A., Kollmann, D., Nöcker, T., Kühling, J., Westerwelle, A., 2018. Sondergutachten 80 – Die Buchpreisbindung in einem sich änderndem Marktumfeld. Technical Report. German Monopolies Commission.

WHO Collaborating Centre for Drug Statistics Methodology, 2024. Guidelines for ATC Classification and DDD Assignment. Oslo. URL: https://atcddd.fhi.no/atc_ddd_index_and_guidelines/guidelines/. last updated: 2024-03-01.

Wong, B.L.H., Maaß, L., Vodden, A., van Kessel, R., Sorbello, S., Buttigieg, S., Odone, A., 2022. The dawn of digital public health in europe: Implications for public health policy and practice. The Lancet Regional Health–Europe 14.

Yeung, K., Basu, A., Hansen, R.N., Sullivan, S.D., 2018. Price elasticities of pharmaceuticals in a value based-formulary setting. Health economics 27, 1788–1804.

Yurkovich, J.T., Evans, S.J., Rappaport, N., Boore, J.L., Lovejoy, J.C., Price, N.D., Hood, L.E., 2024. The transition from genomics to phenomics in personalized population health. Nature Reviews Genetics 25, 286–302.

A Appendix to Section "Methods"

A.1 Reimbursement for Prescription Drugs in Germany

In Germany, most residents are mandated to hold health insurance (Federal Ministry of Health, `https://t.ly/8Ev6L`). This system offers two primary options: statutory health insurance ("Gesetzliche Krankenversicherung", GKV) and private health insurance ("Private Krankenversicherung", PKV). As of 2021, around 73.3 Mio. and 8.7 Mio. citizens were members of the statutory and private health insurance scheme, respectively (Statista, `https://t.ly/pyzwt`). Rx drugs are prescribed by physicians and dispensed by both traditional and online pharmacies. This holds irrespective of the patients' insurance. Reimbursement, however, differs between statutory and private health insurance.

Members of the statutory health insurance are typically required to make co-payments, contributing a portion of the drug's cost. These co-payments are calculated based on the AVP. If the AVP is ...

- ... below EUR 5 the co-payment equals the AVP,
- ... between EUR 5 and EUR 50 the co-payment equals EUR 5,
- ... between EUR 50 and EUR 100 the co-payment equals 10% of the AVP,
- ... above EUR 100 the co-payment is capped at EUR 10.

Further payments beyond the aforementioned co-payments are possible, which depend on rebates between the consumers' insurance company and the drug manufacturers. These rebates fluctuate frequently, often on a quarterly basis, and systematic data on their exact amounts is generally unavailable. The remaining difference between the price of a prescribed drug and the co-payment and any applicable rebates is typically covered by the insurance provider.

Privately insured individuals can select from a range of insurance contracts, each offering distinct reimbursement structures. Typically, these schemes require patients to initially cover costs out-of-pocket, followed by reimbursement from the insurance provider upon submission of invoices. Contracts can be customized to accommodate individual needs, such as making exceptions for high-cost hospitalizations.

A.2 Data Handling and Preprocessing

The data were obtained from three major German suppliers of merchandise information systems (MIS): AWINTA (`https://www.awinta.de`), ADG (`https://www.adg.de`) and Pharmatechnik (`https://www.pharmatechnik.de`). These MIS oversee the entire system of inventory management and provide both hardware and software solutions to pharmacies, essentially handling the IT infrastructure. As a result, our data includes each transaction conducted by a pharmacist with a customer, excluding specialty drugs (e.g. cytostatics).

The dataset encompasses transactions of approximately 9,231 offline pharmacies, covering nearly half of all pharmacies in Germany. The data capture individual sales transactions from January 1, 2018, to October 31, 2022.

Further, essential details like AVP, transaction revenue, patient co-payments, and the central pharmaceutical number (PCN) that uniquely identifies each product are available or can be concluded from the data. Anonymized pharmacy identifiers are provided at a 2-digit zip code level. This represents the highest level of granularity attainable without compromising data privacy.

Using AVP, we can also calculate the wholesale price ("Apothekeneinkaufspreis", AEP) for each product, which allows us to compute the total remuneration pharmacies receive from RX-drugs. This calculation provides the basis for assessing the pharmacies' income from regulated drug sales (see Appendix C.2 for more details on the remuneration calculation).

To account for variations in package sizes, we use the German N-classification system established by the Federal Institute for Drugs and Medical Devices (`https://www.bfarm.de/DE/Arzneimittel/Arzneimittelinformationen/Packungsgroessen/_node.html`). This system categorizes pharmaceutical packages into three size classes based on estimated daily doses: N1 (approximately 10 days), N2 (approximately 30 days), and N3 (approximately 100 days). Given the limited availability of the usual "daily defined doses" (DDD) scheme for the PCN-level in our data set (see WHO Collaborating Centre for Drug Statistics Methodology (2024) and `https://atcddd.fhi.no/ddd/definition_and_general_considera/` on the DDD-scheme), the N-classification system provides a workable alternative for comparing package sizes. In relation to DDD-scheme, we refer to the normalization based on the N-classification scheme as NDD in this Appendix.

The classifications were obtained from IQVIA, which provides standardized package size information for many PCNs. However, to ensure broader coverage, we also integrated publicly available data from the largest German health insurance *Techniker Krankenkasse* (`https://www.tk.de/resource

`/blob/2058850/3f65533a18b118a9ebcf585ef2830c40/rabattvertraege-pzn-liste-gesamt-data.pdf`).

Data on the evolution of health insurance memberships are obtained from the Federal Ministry of Health (`https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/mitglieder-und-versicherte`) and the Association of Private Health Insurance (`https://www.pkv-zahlenportal.de/werte/2012/2022/12/pers-kkv/basket/result`).

Our analysis utilizes annual data aggregated into two datasets: one at the 2-digit zip code level and another at the individual pharmacy level. Data for 2022 was excluded due to incompleteness. We focused on data for prescription drugs (excluding COVID-19 vaccines, prescribed masks, and specialty medications) pre-filtered by MIS suppliers. Using the dictionary *Gelbe Liste Pharmaindex* (`https://www.gelbe-liste.de/`), we further excluded non-pharmaceutical items like medical devices. The analysis only considers direct drug dispenses to customers, excluding sales via courier services and to nursing homes.

To ensure data consistency, we limited the analysis to standardized package sizes (N1, N2, and N3 categories) mandated by statutory health insurance. We balanced and pruned the data to exclude inactive pharmacies, those with minimal sales, or those specializing in expensive medications. Additionally, to minimize the influence of outliers, we removed the top and bottom 2.5% of pharmacies based on sales and remuneration distributions

B Appendix to Section "Results"

B.1 Descriptive Statistics

This section provides more detailed information on the geographical coverage of our sample, a graphical analysis of the DiD effect and the usual summary statistics.

Figure B.1 visually depicts the geographic scope of pharmacy sales data across Germany. The left panel illustrates the overall sample coverage, encompassing 9,231 unique pharmacies out of a total of 18,839 (49 %). The right panel zooms in on the coverage of the final sample, which comprises 5,487 pharmacies (29 %). Color intensity within each region corresponds to the percentage of coverage, with darker shades signifying higher data representation. A comparison of both panels reveals that, despite data balancing and processing, our sample retains coverage across all 2-digit zip codes in Germany, with a median coverage of approximately 29 % per 2-digit zip code.

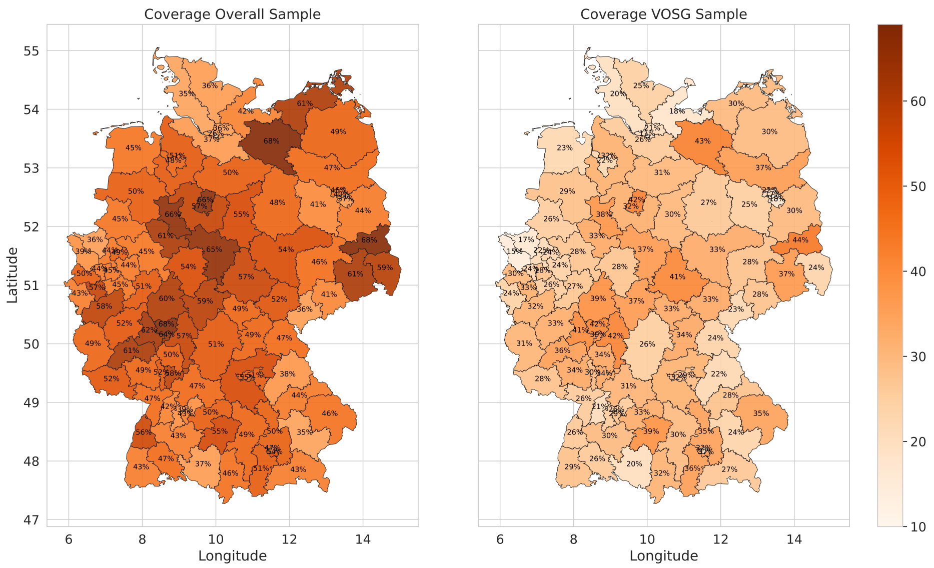


Figure B.1: Coverage of Pharmacy Sales Data Across German Regions. Source: MIS suppliers and web page of *Apothekenumschau* (`https://www.apotheken-umschau.de/apothekenfinder/`), scraped on September 29, 2020.

Figure B.2 depicts the annual sales trends in packages and NDD (see Section A.2), categorized by statutory health and private prescriptions. It comprises four subfigures: the top ones show absolute values, while the bottom ones display sales normalized to 2020. Figure B.2 reveals two key points.

Based on the figure, one can see that prescriptions issued to members of statutory health insurance are six times higher than those issued to privately insured and self-pay patients. This discrepancy is explained by the fact that most Germans are covered by statutory health insurance (see Appendix A.1).

We can also calculate a DiD effect by aggregating our data in years and by statutory health and private prescriptions without any covariates. That way, we find an effect of the VOASG in packages (NDD) of approximately 0.0163 (0.0113). Accordingly, by comparing means we find that sales of offline channels increased by 1.63 % (1.13 %) due to VOASG, compared to a counterfactual of a state without VOASG. This result should be viewed as a first indication. It is refined using a more sophisticated approach (see Section 3 and B.2).

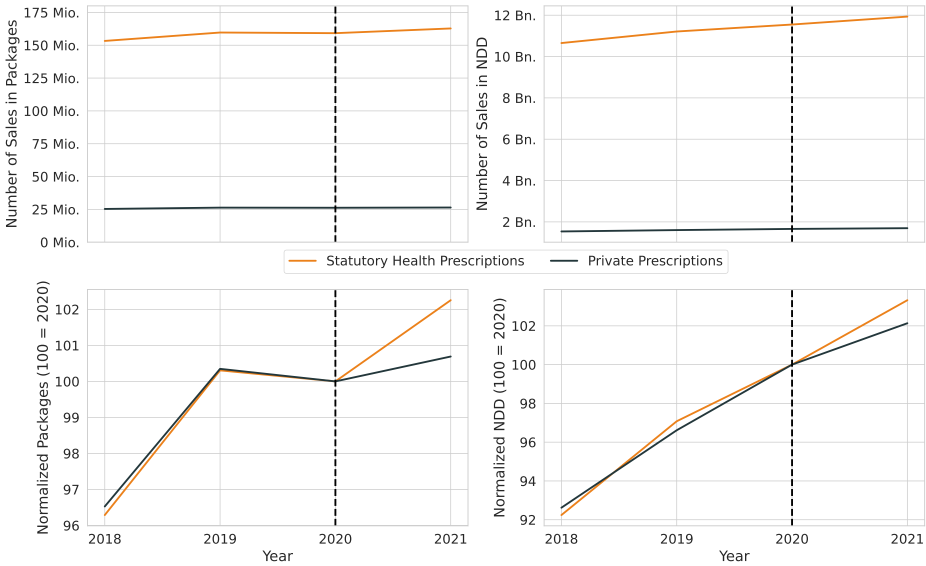


Figure B.2: Evolution of sales in packages and NDD by statutory health and private prescriptions. The top figures denote sales in absolute values, while the bottom figures depict sales normalized to the year 2020.

Tables B.1 and B.2 report summary statistics at the 2-digit zip code and pharmacy levels, respectively, differentiated between private and statutory health prescriptions.

	Private Prescriptions						Statutory Health Prescriptions					
	N	Mean	SD	Min	Median	Max	N	Mean	SD	Min	Median	Max
Sales per 2-digit-zip code and Year in Packages	380	274,393	126,162	33,328	274,106	709,750	380	1,670,946	647,546	177,155	1,594,334	3,948,482
Sales per 2-digit-zip code and Year in NDD	380	17,088,357	7,877,312	1,958,510	16,797,645	46,334,170	380	1.19e+08	47,241,362	11,561,710	1.16e+08	3.06e+08
Gross Revenue per 2-digit-zip code and Year in Euro	380	14,668,064	6,949,777	1,775,727	14,312,372	39,087,452	380	93,061,340	36,210,488	11,564,190	92,083,137	2.39e+08
Net Revenue per 2-digit-zip code and Year in Euro	380	12,366,244	5,861,491	1,492,210	12,104,528	32,847,140	380	78,456,469	30,537,605	9,717,812	77,561,420	2.01e+08
Net Revenue (w.o. lump sum fees) per 2-digit-zip code and Year in Euro	380	12,314,791	5,837,464	1,486,483	12,050,517	32,691,821	380	75,653,612	29,463,615	9,425,507	7.5e+07	1.94e+08
Total Remuneration per 2-digit-zip code and Year in Euro	380	2,583,160	1,191,630	315,717	2,574,276	6,706,106	380	13,332,543	5,159,417	1,454,439	12,815,406	31,943,711
Average Remuneration per 2-digit-zip code and Year in Euro	380	9.4	0.084	9.15	9.41	9.71	380	7.98	0.0927	7.77	7.97	8.33
Gross AVP per 2-digit-zip code and Year in Euro/Package	380	53	3.38	42.7	53.1	65.7	380	55.9	3.75	47.2	55.5	69.3
Net AVP per 2-digit-zip code and Year in Euro/Package	380	44.7	2.89	35.9	44.8	55.2	380	47.1	3.19	39.7	46.8	59
Number of Pharmacies per 2 digit zip-code	380	57.8	21.8	9	54	121	380	57.8	21.8	9	54	121
Quantity weighted average NDD	380	62.2	2.61	54.2	62.4	68.2	380	71.3	3.06	64	71	78.8

Table B.1: Summary statistics on the 2-digit-zip code sample. 760 observations in total.

	Private Prescriptions						Statutory Health Prescriptions					
	N	Mean	SD	Min	Median	Max	N	Mean	SD	Min	Median	Max
Sales per Pharmacy and Year in Packages	21948	4,751	2,685	385	4,162	25,207	21948	28,930	12,278	6,591	26,688	74,147
Sales per Pharmacy and Year in NDD	21948	295,862	167,431	22,640	259,715	1,530,590	21948	2,066,169	912,123	385,530	1,907,845	5,777,330
Gross Revenue per Pharmacy and Year in Euro	21948	253,958	170,740	12,770	213,515	2,359,418	21948	1,611,232	763,634	313,249	1,458,762	7,044,259
Net Revenue per Pharmacy and Year in Euro	21948	214,105	143,972	10,731	180,068	1,982,707	21948	1,358,368	643,953	263,235	1,228,696	5,919,541
Net Revenue (w.o. lump sum fees) per Pharmacy and Year in Euro	21948	213,214	143,510	10,636	179,253	1,981,299	21948	1,309,840	625,601	251,675	1,183,001	5,820,396
Total Remuneration per Pharmacy and Year in Euro	21948	44,724	25,607	3,605	39,117	242,996	21948	230,835	98,256	51,762	212,854	594,033
Average Remuneration per Pharmacy and Year in Euro	21948	9.39	0.385	8.62	9.32	24.1	21948	7.98	0.279	7.41	7.92	9.84
Gross AVP per Pharmacy and Year in Euro/Package	21948	52.4	15.7	21.3	49.7	652	21948	55.8	11.3	32.5	53.4	132
Net AVP per Pharmacy and Year in Euro/Package	21948	44.2	13.2	17.9	41.9	548	21948	47.1	9.57	27.3	45	111
Quantity weighted average NDD	21948	62.3	5.16	30	62.8	78.7	21948	71.1	5.64	36	71.7	85

Table B.2: Summary statistics on the pharmacy level sample. 43,896 observations in total.

The descriptive statistics include the number of observations (N), mean, standard deviation (SD), median, minimum, and maximum for annual sales in packages or NDD, gross revenue, net revenue, net revenue excluding lump sum fees, total remuneration, quantity weighted average NDD, average remuneration, and both gross and net AVP, presented separately for 2-digit zip codes and pharmacies. These sales figures represent the main variables of interest in our estimations, while the other variables are used to calculate distributional effects in Section 4.1.

B.2 DiD-Estimation

The causal effect of the VOASG is examined through a TWFE DiD estimation in conjunction with an event study. The empirical model we estimate is given by

$$\ln(y_{pgt}) = \alpha_{pg} + \gamma_t + \beta D_{pgt} + W_{pgt}\mu + \epsilon_{pgt}, \tag{B.1}$$

where y denotes sales for pharmacies p and statutory health or private prescriptions g at time t (in years). The outcome variable y is measured in logs. D_{pgt} is an indicator variable that equals one if g represents a statutory health prescription at time $t = 2021$. Fixed effects α_{pg} and γ_t are included to capture cross-sectional heterogeneity for each combination of p and g, as well as time-varying effects. Therefore, β represents the DiD effect or ATT. Since the outcome variable is measured in logs, the coefficient β can be interpreted as a percentage change. The exact ATT is computed by $e^{\beta} - 1$, although this transformation is negligible when the effect is close to zero, as it is in our case.

The matrix $\boldsymbol{W}_{pgt}$ includes the following covariates: (i) the quantity-weighted average of doses (NDD) per pharmacy p and group g at time t, and (ii) the fraction of members in statutory health or private insurance at the national level. Covariate (i) accounts for potential temporal variations in package sizes between treatment and control groups. Covariate (ii) controls for potential changes in group sizes and technically has a similar effect as a time trend. Given collinearity between Covariate (ii) and a direct time trend, we cannot include both in a single model.

The results of the TWFE DiD estimation (B.1) are presented in Section 3, Table 1. The ATT varies between 0.0136 and 0.0165.

Alternatively, we can use NDD as the dependent variable in equation (B.1) instead of incorporating it as a covariate. This approach provides a robustness check for the estimates presented earlier. Table B.3 displays the results of this alternative estimation.

The table's interpretation and structure mirror those of Table 1. The ATT reported in Table B.3 are comparable to those in the main text, ranging from 0.0139 to 0.0158.

B.3 Event Study

The estimation equation for the event study presented in Section 3 reads as follows:

$$\ln(y_{pgt}) = \alpha_{pg} + \gamma_t + \sum_{\tau=2018}^{2020} \delta_\tau D_{pg\tau} + \beta_{2021} D_{pg2021} + W_{pgt}\mu + \epsilon_{pgt}. \qquad \text{(B.2)}$$

The model includes the years 2018, 2019 and 2020 as leads and the year 2021 as a lag, denoted by δ_{2018}, δ_{2019}, δ_{2020}, and β_{2021}, respectively. As is standard in the literature (see Cunningham (2021, Chapter 9.4) or Freyalden-

	(1)	(2)	(3)	(4)
Specification A: Sales in NDD				
DiD-Coefficient	0.0139***	0.0150***	0.0150***	0.0150***
	(0.0038)	(0.0023)	(0.0015)	(0.0031)
Fraction of Members in Insurance	-1.9188	-2.1932*	-2.1932*	-2.1932
	(1.8371)	(1.2062)	(1.1068)	(1.6978)
Weighted Average of Doses	0.0172***	0.0156***	0.0156***	0.0156***
	(0.0029)	(0.0006)	(0.0008)	(0.0008)
Observations	760	43,896	43,896	43,896
Adj. R2	0.9998	0.9945	0.9945	0.9945
FE: Year	X	X	X	X
FE: 2 digit zip code & Treated	X			
FE: Pharmacy & Treated		X	X	X
Std. Errors	Robust	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated
Specification B: Sales in NDD with Trends				
DiD-Coefficient	0.0145***	0.0158***	0.0158***	0.0158***
	(0.0047)	(0.0028)	(0.0018)	(0.0036)
Weighted Average of Doses	0.0171***	0.0156***	0.0156***	0.0156***
	(0.0030)	(0.0006)	(0.0008)	(0.0008)
Time Trend of Treatment-Group	-0.0013	-0.0015	-0.0015	-0.0015
	(0.0017)	(0.0011)	(0.0010)	(0.0015)
Observations	760	43,896	43,896	43,896
Adj. R2	0.9998	0.9945	0.9945	0.9945
FE: Year	X	X	X	X
FE: 2 digit zip code & Treated	X			
FE: Pharmacy & Treated		X	X	X
Std. Errors	Robust	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated

Note:
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: TWFE DiD estimation with NDD as dependent variable.

hoven et al. (2019)), the event study is normalized to the pre-intervention period (2020). (Note that it is technically impossible to include a time trend or Covariate (ii) as introduced in Section B.2 into an event study as it would be collinear with the leads and lags.) The coefficient β_{2021} in Equation (B.2) can be interpreted in a similar way as the ATT, because there is only one post-treatment period. The coefficients $\delta_\tau \in \{2018, 2019, 2020\}$ capture the pre-treatment effects, which allow us to examine dynamics prior to the policy intervention.

The pre-treatment coefficients, δ_{2018} and δ_{2019}, can be used to assess the parallel trends assumption before the treatment. Based on the literature (Roth et al., 2023, Sections 4.3–4.4), parallel trends are plausible when pre-treatment coefficients are statistically insignificant. Figure 1 in the main Text and Table B.4 below shows that this holds for all lags.

	(1)	(2)	(3)	(4)
Lead 2018	-0.0035	-0.0026	-0.0026	-0.0026
	(0.0035)	(0.0022)	(0.0021)	(0.0030)
Lead 2019	0.0011	0.0023	0.0023*	0.0023
	(0.0026)	(0.0016)	(0.0012)	(0.0021)
Lag 2021	0.0162***	0.0174***	0.0174***	0.0174***
	(0.0039)	(0.0023)	(0.0014)	(0.0031)
Weighted Average of Doses	0.0009	-0.0006	-0.0006	-0.0006
	(0.0031)	(0.0006)	(0.0009)	(0.0008)
Observations	760	43,896	43,896	43,896
Adj. R2	0.9997	0.9938	0.9938	0.9938
FE: Year	X	X	X	X
FE: 2 digit zip code & Treated	X			
FE: Pharmacy & Treated		X	X	X
Std. Errors	Robust	Robust	by: 2-digit-zip code	by: 2-digit-zip code & Treated

Note:
* p < 0.1, ** p < 0.05, *** p < 0.01

Table B.4: Estimation results for Equation (B.2)).

C Appendix to Section "Discussion"

C.1 DiD with Deciles

To estimate how the effects of the VOASG differ between pharmacies with different revenues, the DiD coefficient D_{gt} is interacted with a dummy variable $Deciles_{pgt}$ that categorizes pharmacies into deciles based on their total revenue:

$$ln(y_{pgt}) = \alpha_{pg} + \gamma_t + Deciles_{pgt} + Deciles_{pgt} \times D_{pgt} + W_{pgt}\mu + \epsilon_{pgt}. \quad \text{(C.1)}$$

Equation (C.1) closely resembles Equation (B.1). The dummy $Deciles_{pgt}$ accounts for decile-specific heterogeneity. (Note that $Deciles_{pgt}$ is not collinear with the fixed effect α_{pg} as pharmacies can shift between deciles over time). This approach addresses unobserved heterogeneity.

C.2 Remuneration

In this section, we describe how pharmacies' remuneration is computed. This is a prerequisite step to quantify the absolute gains caused by the introduction of VOASG in the following section.

The remuneration structure for pharmacies dispensing prescription medications in Germany is subject to stringent regulatory oversight. Pharmacists are currently compensated with a fixed fee of € 8.35 per package, in addition to a variable component that constitutes 3 % of the AEP (§ 3 AMPreisV). For prescriptions covered by statutory health insurance, pharmacies are required to deduct an additional gross lump-sum fee of € 1.77 (§ 130 (1) SGB V), provided the insurance pays within 10 days. In net terms, this results in a deduction of € 1.49 per package.

This remuneration framework applies exclusively to *Fertigarzneimittel* (finished dosage form, § 4 (1) sentence 2 German Medicines Act, AMG), which is the only category included in our dataset. The following computation, based on AVP, will clarify how the remuneration is calculated, as our analysis focuses solely on AVP for prescription drugs.

The remuneration for pharmacies can be computed based on AVP, as it represents a list price that follows a consistent pattern:

$$(Gross)\ AVP = (1 + \text{VAT}) \cdot (Net)\ AVP \tag{C.2}$$

$$(Net)\ AVP = r_f + r_v + pDL + nDZ + AEP. \tag{C.3}$$

As shown in Equation (C.2), gross AVP is calculated by multiplying net AVP by the value-added tax rate, which is 19% in Germany (16% for the first two quarters of 2020). Net AVP, Equation (C.3), consists of a fixed rate, $r_f =$ € 8.35, in addition to a variable component, $r_v = 0.03 \cdot AEP$. Recall that AEP refers to the wholesale price. Moreover, AVP includes two lump-sum fees per prescription drug in net terms: the *pharmazeutische Dienstleistung* (pDL) fee, $pDL =$ € 0.20, and the *Notdienstzuschlag* (nDZ) fee, $nDZ =$ € 0.20 (for more details, refer to `https://www.abda.de/apotheke-in-deutschland/preise-und-honorare/beispielrechnung/`). By rearranging the terms of Equation (C.3), we can derive AEP:

$$AEP = \frac{(Net)\ AVP - (r_f + pDL + nDZ)}{1.03}. \tag{C.4}$$

Based on AEP, we can compute the compensation per Rx-drug dispensed:

$$\begin{aligned} R(AEP) &= \underbrace{(Net)\ AVP - pDL - nDZ - STHF}_{(Net)\ AVP \text{ w.o. lump sum fees}} - AEP \\ &= r_f + r_v - STHF \\ &= € \ 8.35 + 0.03 \cdot AEP - € \ 1.49. \end{aligned} \tag{C.5}$$

Here, $STHF =$ € 1.49 denotes the additional net lump-sum fee per statutory health insurance prescription drug (with $STHF = 0$ for private prescriptions). For example, an Rx-drug prescribed under statutory health insurance with $AEP =$ € 50 results in a remuneration of $R(50) =$ € 8.36.

Regarding the technical implementation, during the aggregation process, we ascertain net revenue excluding elements pDL, nDZ, and $STHF$, denoted $(Net)\ AVP\ w.o.\ lump\ sum\ fees$. We also determine the total costs by summing over AEP and further differentiate these figures to derive the total remuneration for each 2-digit zip code or pharmacy p, group g, and year t (see Tables B.1 and B.2). The results of the respective calculations for deciles 1 to 10 are summarized in Table C.1.

	Deciles	Mean	SD	Min	P25	Median	P75	Max
Gross Revenue (in Euro) from Statutory Health Prescriptions by Pharmacy in 2021	D01	694,825	107,027	365,910	619,466	707,555	780,411	918,617
	D02	898,429	105,612	375,124	836,078	907,012	978,202	1,118,446
	D03	1,069,323	105,645	695,538	1,008,279	1,082,769	1,144,722	1,304,631
	D04	1,217,061	116,976	519,110	1,153,307	1,229,121	1,298,209	1,447,576
	D05	1,377,001	133,080	790,864	1,307,640	1,397,290	1,462,790	1,642,521
	D06	1,559,251	142,510	834,556	1,487,207	1,578,656	1,659,385	1,849,085
	D07	1,763,678	147,199	732,421	1,676,696	1,771,491	1,856,267	2,074,630
	D08	2,011,384	175,497	1,323,046	1,914,402	2,025,218	2,133,129	2,416,269
	D09	2,381,734	233,485	1,006,973	2,243,196	2,399,030	2,541,571	2,945,108
	D10	3,249,097	627,266	2,012,603	2,854,133	3,098,203	3,482,232	6,869,854
Net Revenue (in Euro) from Statutory Health Prescriptions by Pharmacy in 2021	D01	583,871	89,935	307,466	520,558	594,562	655,778	771,945
	D02	754,962	88,741	315,225	702,584	762,193	821,996	939,868
	D03	898,569	88,772	584,445	847,291	909,839	961,945	1,096,164
	D04	1,022,717	98,294	436,225	969,085	1,032,873	1,090,928	1,216,397
	D05	1,157,115	111,837	664,572	1,098,846	1,174,192	1,229,234	1,380,185
	D06	1,310,257	119,747	701,306	1,249,705	1,326,587	1,394,440	1,553,846
	D07	1,482,054	123,684	615,481	1,408,982	1,488,629	1,559,907	1,743,322
	D08	1,690,203	147,481	1,111,802	1,608,663	1,701,839	1,792,340	2,030,473
	D09	$2e+06$	196,206	846,196	1,885,031	2,015,969	2,135,773	2,474,939
	D10	2,730,298	527,110	1,691,216	2,398,426	2,603,527	2,926,157	5,772,980
Net Revenue w.o. lump sum fees (in Euro) from Statutory Health Prescriptions by Pharmacy in 2021	D01	560,415	87,060	295,819	498,801	568,880	628,475	740,249
	D02	725,822	85,601	301,853	677,522	732,485	789,090	906,640
	D03	864,615	85,838	556,555	814,771	875,090	926,494	1,060,352
	D04	984,163	95,004	420,673	932,825	992,887	1,051,264	1,175,963
	D05	1,114,338	108,121	646,402	1,056,123	1,130,154	1,183,478	1,347,118
	D06	1,262,126	115,626	677,613	1,202,639	1,277,625	1,339,810	1,504,633
	D07	1,428,648	120,700	597,998	1,357,673	1,435,713	1,507,907	1,689,007
	D08	1,629,473	142,867	1,070,758	1,547,803	1,643,388	1,729,281	1,966,905
	D09	1,930,876	189,985	826,700	1,817,743	1,944,083	2,058,881	2,393,400
	D10	2,644,192	520,429	1,630,590	2,316,601	2,516,686	2,841,886	5,664,327
Total Remuneration (in Euro) from Statutory Health Prescriptions by Pharmacy in 2021	D01	107,788	17,536	54,039	95,022	107,500	120,585	153,070
	D02	134,769	22,370	60,962	120,310	135,985	150,501	187,565
	D03	157,586	23,512	93,360	141,666	160,836	174,790	217,194
	D04	179,000	26,270	72,882	162,217	180,972	199,126	233,019
	D05	199,269	31,633	89,633	178,379	204,688	221,627	270,290
	D06	224,442	35,606	112,155	201,281	229,664	250,671	303,367
	D07	249,862	36,033	85,547	229,440	250,915	274,787	329,040
	D08	284,276	43,145	139,039	258,801	288,453	314,700	397,559
	D09	331,314	55,485	$1e+05$	298,599	335,423	372,308	470,201
	D10	412,801	75,463	207,936	367,546	412,764	465,584	588,611
Total Net Revenue (in Euro) per Pharmacy from Rx in 2021	D01	693,373	84,438	436,844	631,480	713,541	763,667	805,599
	D02	893,718	48,541	805,975	851,774	898,746	934,183	973,616
	D03	1,051,556	44,295	974,228	1,013,446	1,053,345	1,092,117	1,123,415
	D04	1,193,845	42,963	1,123,768	1,154,162	1,193,204	1,230,982	1,271,731
	D05	1,351,620	44,027	1,272,385	1,313,041	1,350,629	1,390,098	1,427,260
	D06	1,512,879	48,487	1,427,668	1,470,210	1,514,149	1,556,733	1,594,732
	D07	1,703,807	61,527	1,595,295	1,652,130	1,702,838	1,759,689	1,808,252
	D08	1,938,898	80,837	1,808,417	1,869,584	1,934,012	2,009,602	2,088,366
	D09	2,306,836	132,220	2,088,832	2,197,524	2,297,658	2,419,661	2,556,098
	D10	3,108,563	538,309	2,556,484	2,729,172	2,939,959	3,332,300	6,242,840
Sales (in Packages) of Statutory Health Prescriptions per Pharmacy in 2021	D01	13,733	2,350	6,820	11,952	13,624	15,402	19,821
	D02	17,060	3,132	7,833	14,927	17,239	19,288	24,610
	D03	19,879	3,330	10,348	17,666	20,220	22,196	28,623
	D04	22,572	3,723	9,103	20,151	22,780	25,439	30,227
	D05	25,046	4,516	10,631	21,999	25,777	28,296	35,292
	D06	28,179	5,102	13,402	24,852	28,706	31,961	39,949
	D07	31,267	5,227	10,229	28,200	31,482	34,974	42,797
	D08	35,556	6,199	15,360	31,834	36,198	39,817	52,179
	D09	41,300	7,939	11,428	36,436	41,830	46,890	61,203
	D10	50,415	10,298	23,148	43,717	51,290	57,808	73,556
Average Remuneration per Sale (in Euro) from Statutory Health Prescriptions by Pharmacy in 2021	D01	7.86	0.157	7.52	7.75	7.84	7.96	8.42
	D02	7.93	0.217	7.54	7.78	7.88	8.04	8.74
	D03	7.96	0.215	7.57	7.82	7.93	8.06	9.07
	D04	7.96	0.209	7.57	7.82	7.92	8.05	8.97
	D05	7.99	0.257	7.61	7.84	7.93	8.11	9.32
	D06	8	0.258	7.55	7.84	7.95	8.1	9.68
	D07	8.03	0.273	7.61	7.86	7.97	8.12	9.65
	D08	8.03	0.271	7.62	7.86	7.97	8.11	9.35
	D09	8.07	0.307	7.61	7.88	7.99	8.17	9.53
	D10	8.24	0.391	7.66	7.96	8.11	8.42	9.84

Table C.1: Summary statistics for the presented calculations by deciles 1 to 10.

C.3 Distributional Effects

In Section 4.1, particularly in Figure 2, we provided a histogram visualizing the "Average Gain from VOASG" for each pharmacy. This measure is derived from the DiD coefficient *for each decile*, the volume of sales from statutory health prescriptions, and the average remuneration per sale across pharmacies for the year 2021 (see Table C.1). The average remuneration per sale is calculated by dividing the total remuneration by the sales expressed in packages, thereby yielding a quantity-weighted average remuneration per package for each pharmacy p, group g, and year t.

Table C.2 displays summary statistics for the *Average gain from VOASG* by decile.

	Deciles	Mean	SD	Min	P25	Median	P75	Max
Average Gain from VOASG (in Euro) by $\beta_{Deciles}$	D01	1,359	221	681	1,198	1,355	1,520	1,929
	D02	1,484	246	671	1,325	1,498	1,657	2,066
	D03	1,671	249	990	1,502	1,706	1,854	2,303
	D04	2,569	377	1,046	2,328	2,598	2,858	3,345
	D05	2,508	398	1,128	2,245	2,576	2,789	3,401
	D06	2,053	326	1,026	1,841	2,101	2,293	2,775
	D07	3,236	467	1,108	2,971	3,249	3,558	4,261
	D08	5,072	770	2,481	4,618	5,147	5,615	7,094
	D09	6,761	1,132	2,045	6,093	6,845	7,597	9,595
	D10	7,688	1,405	3,872	6,845	7,687	8,671	10,962

Table C.2: Summary statistics for the *Average gain from VOASG* by decile.

5. Conclusion

This doctoral thesis examines how regulatory changes, market structures, and government intervention shape outcomes in transport, retail, and health markets. Across seven manuscripts, the research focuses on the effects of structural changes and policy interventions, using empirical methods such as Data Envelopment Analysis (DEA), Fixed-Effects Models, Difference-in-Differences (DID), and Instrumental Variable techniques.

The first manuscript analyzes railway markets in Europe, focusing on how government funding structures influence efficiency. The results suggest that allocating a larger share of public contributions upstream to infrastructure managers can improve network utilization, potentially through mechanisms such as lower access charges and increased competition.

The second manuscript studies the interurban bus market in Germany. It shows how a dominant firm, Flixbus, used strategic brand proliferation to deter competition during a period of market transition. The findings also highlight the interdependence of bus and rail markets, where competition in one mode affects outcomes in the other.

The retail-focused manuscripts examine strategic behavior in grocery and book markets. The study of Aldi's shopping hours demonstrates how firms use extended operating hours as a strategic tool to enter new market segments, such as late-night shopping, while adapting to local competitive conditions. The analysis of the book market reveals that physical bookstores and e-commerce are imperfect substitutes, with regulatory frameworks, such as fixed book prices, playing a significant role in shaping market outcomes.

The fifth manuscript analyzes the efficiency of COVID-19 vaccination campaigns across Germany's federal states. The results reveal significant regional differences and show how expanding distribution networks, such as integrating doctors' offices, improved efficiency. These findings emphasize the role of organization and resource allocation in achieving effective public health outcomes.

The last manuscript evaluates the regulation of Rx rebates in Germany. It investigates how the 2020 Local Pharmacy Support Act affected competition and profitability in the pharmacy market. The findings show that while the policy modestly increased offline pharmacy sales, it disproportionately benefited high-revenue pharmacies and failed to reverse the long-term decline in the number of brick-and-mortar pharmacies.

Overall, this thesis provides insights into how regulatory changes and market structures shape economic outcomes across sectors. The manuscripts demonstrate the importance of understanding the incentives of firms, consumers, and public agencies to improve efficiency and competitiveness. Policymakers and industry participants can draw on these findings to design better policies and strategies tailored to specific market conditions. Future research could extend these analyses to other sectors or regions, addressing new challenges as markets evolve.

References

Agomo, C. O. (2012). The role of community pharmacists in public health: A scoping review of the literature. *Journal of Pharmaceutical Health Services Research*, *3*(1), 25–33.

Beestermöller, M. (2017). Striking Evidence? Demand Persistence for Inter-City Buses from German Railway Strikes. *Münchener wirtschaftswissenschaftliche Beiträge*, *2*.

Böckers, V., Haucap, J., Heimeshoff, U., & Thorwarth, S. (2015). Auswirkungen der Fernbusliberalisierung auf den Schienenpersonenfernverkehr. *List Forum für Wirtschafts-und Finanzpolitik*, *41*(1), 75–90.

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.

Cantos, P., Pastor, J. M., & Serrano, L. (1999). Productivity, efficiency and technical change in the european railways: A non-parametric approach. *Transportation*, *26*, 337–357.

Cavallo, A. (2017). Are online and offline prices similar? evidence from large multi-channel retailers. *American Economic Review*, *107*(1), 283–303.

Dixit, A. (1982). Recent developments in oligopoly theory. *The American Economic Review*, *72*(2), 12–17.

ECDC. (2021). *Rollout of covid-19 vaccines in the eu/eea: Challenges and good practice* (tech. rep.) (https://www.ecdc.europa.eu/en/publications-data/rollout-covid-19-vaccines-eueea-challenges-and-good-practice). European Centre for Disease Prevention and Control.

Ellison, G., & Ellison, S. F. (2011). Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration. *American Economic Journal: Microeconomics*, *3*(1), 1–36.

Flores, M., & Wenzel, T. (2016). Shopping hours and price competition with loyal consumers. *The BE Journal of Economic Analysis & Policy*, *16*(1), 393–407.

Friederiszick, H. W., Röller, L.-H., & Schulz, C. C. (2003). Evaluation of the effectiveness of state aid as a policy instrument: The Railway Sector. *Draft Version, 24.10.2003*.

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

Goolsbee, A., & Syverson, C. (2008). How do incumbents respond to the threat of entry? evidence from the major airlines. *The Quarterly journal of economics*, *123*(4), 1611–1633.

Götz, G., & Schäfer, J. T. (2020). Financing railways. In *Handbook on railway regulation* (pp. 304–321). Edward Elgar Publishing.

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.

Inderst, R., & Irmen, A. (2005). Shopping hours and price competition. *European Economic Review*, *49*, 1105–1124.

Kosfeld, M. (2002). Why shops close again: An evolutionary perspective on the deregulation of shopping hours. *European Economic Review*, *46*, 51–72.

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

Mills, M. C., & Salisbury, D. (2021). The challenges of distributing covid-19 vaccinations. *EClinicalMedicine*, *31*.

Oum, T. H., & Yu, C. (1994). Economic Efficiency of railways and Implications for Public Policy: A comparative study of the OECD countries. *Journal of Transport Economics and Policy*, 121–138.

Ratzan, S., Schneider, E. C., Hatch, H., & Cacchione, J. (2021). Missing the point—how primary care can overcome covid-19 vaccine “hesitancy”. *New England Journal of Medicine*.

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.

Schäfer, J. T., & Götz, G. (2017). Public Budget Contributions to the European Rail Sector. *Review of Network Economics*, *16*(2), 89–123. https://doi.org/10.1515/rne-2017-0044

Schmalensee, R. (1978). Entry deterrence in the ready-to-eat breakfast cereal industry. *The Bell Journal of Economics*, 305–327.

Shy, O., & Stenbacka, R. (2006). Service hours with asymmetric distributions of ideal service time. *The International Journal of Industrial Organization*, *24*(4), 764–771.

Simar, L., & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of econometrics*, *136*(1), 31–64.

Smith, F. (2009). Private local pharmacies in low-and middle-income countries: A review of interventions to enhance their role in public health. *Tropical Medicine & International Health*, *14*(3), 362–372.

Staat, M. (2006). Efficiency of hospitals in germany: A dea-bootstrap approach. *Applied Economics*, *38*(19), 2255–2263.

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

Toivanen, O., & Waterson, M. (2005). Market structure and entry: Where's the beef. *The RAND Journal of Economics*, *36*(3), 680–699.

Torres, I., Lopez-Cevallos, D., Artaza, O., Profeta, B., Kang, J., & Machado, C. V. (2021). Vaccine scarcity in lmics is a failure of global solidarity and multi-lateral instruments. *The Lancet*, *397*(10287), 1804.

Valdmanis, V., Walker, D., & Fox-Rushby, J. (2003). Are vaccination sites in bangladesh scale efficient? *International Journal of Technology Assessment in Health Care*, *19*(4), 692–697.

A. Other Manuscripts

A.1. Public budget contributions to the European rail sector

Authors: Jan Thomas Schäfer and Georg Götz

Published in: *Review of Network Economics*, vol. 16, no. 2, 2017, pp. 89-123.

DE GRUYTER Rev Netw Econ 2017; 16(2): 89–123

Jan Thomas Schäfer and Georg Götz*

Public Budget Contributions to the European Rail Sector

An in Depth Analysis for Eight Countries

https://doi.org/10.1515/rne-2017-0044

Abstract: This paper provides an analysis of the funding structure of the railways in eight European countries. It updates and expands the well-known database on public contributions to rail, which has initially been published by [NERA (2004) Study of the Financing of and Public Budget Contributions to Railways: A Final Report for European Commission, DG TREN. London: NERA National Economic Research Associates.]. The analysis shows that there are large differences concerning the focus of granted funds which can be explained by different policy objectives, differences in the level and degree of network access charges and different cost coverage ratios of public transport services. We identify a tendency towards two main financing models. In our data-set countries either focus their support payments on the operation of the infrastructure, which implies lower network charges and thus a lower amount of necessary Public Service Compensations, or they focus on the support of transport services with a higher degree of cost coverage of network charges and thus a lower amount of operating contributions paid to the infrastructure manager. The structure of funds, different approaches of infrastructure financing and differences in the treatment of historical debt are likely to have an influence on the performance of the investigated railway systems.

Keywords: contributions; Europe; financing; public budget; railways; subsidies.

JEL Classification: R42; R48; L92; H54.

***Corresponding author: Georg Götz,** Professor, Justus Liebig University Giessen, Giessen, Germany, e-mail: georg.goetz@wirtschaft.uni-giessen.de. http://orcid.org/0000-0001-7878-6726
Jan Thomas Schäfer: Justus Liebig University Giessen, Giessen, Germany

1 Introduction

Public contributions are one of the main funding sources of the European railway sector. Environmental and social reasons as well as the potential to foster economic development are common justifications for these payments. Even though government expenditures for the railway sector make up a significant share of state budgets in most of European countries, detailed information on the level and use of funds are typically not available in easily accessible, systematic and well-structured form. This is why researchers argue that there is plenty of room for a deeper investigation of public support to the European railway sector (see, e.g. Arrigo and Di Foggia 2014: p. 33). Especially in times of tightening budgets and higher financial needs of railway undertakings the debate on the future design of the European railway sector might benefit from an in depth review of different funding schemes. Apart from budgetary and efficiency considerations, detailed data on the level of public contributions to the different organizational units of the sector might also have regulatory relevance since large parts of regulated access charges are paid from public service compensations.

NERA (2004) is one of the few databases that allow for cross-country comparisons of public support in certain categories and therefore goes beyond the rather aggregate figures that are stated in most analyzes. The prevalence of those aggregate approaches (see, e.g. Nash et al. 2011 and van de Velde et al. 2012) are due to the fact that payments are typically not collected on a comparable basis. Furthermore, European legislation does not provide a clear classification of government support since permitted payments are regulated by several directives and guidelines. Because of this, no standardized methods for the assessment of government support to the railway sector do exist. However, given the rather complex financing structures railways exhibit, and the fact that government support occurs on different stages by the means of different methods and might be complementing or substituting other sources of funds, it is obvious that in depth analysis requires a sufficiently detailed database.

In this paper we provide an analysis of the financing structure of the railways in eight European countries. The main contribution of the study can be seen in an update and extension of the well-known database on public contributions to rail which has been initially published by NERA (2004). The database will also contribute to the debate on the future design of the railway sector in Europe. For this we take a deeper look at the financing structure of each railway sector, collected relevant payments and evaluated the consistency of already existing studies. Data have been mainly collected from public budget plans, annual reports of infrastructure managers and transport undertakings, publications of regulatory authorities and statistical offices. We furthermore conducted a survey to obtain additional information on the level and breakdown of government support. In our data we differ between seven categories of government support that reflect the organizational structure of the European railway sector as well as the main areas of public contributions. Our database covers a period from 2001 to 2015 and includes financial figures as well as other key characteristics for the railway sectors of France, Germany, Great-Britain, Italy, Norway, Spain, Sweden and Switzerland. To be able to compare the level of government support and the development of transfers over the observation period we adjusted our figures using passenger-ton-kilometer (ptkm) and the number of inhabitants (tax payer cost) as the main indicators.

2 Literature Review

2.1 The NERA (2004) Database

The publication of NERA (2004) is one of the most detailed, publicly available, studies of public contributions to the European railway sector. It compares direct public budget contributions of 17 European countries with state aid data from the European Commission. The payments are summarized into several groups, reflecting different types of public budget contributions permitted by EU legislation. This includes compensation for Public Service Obligations, freight transport grants, support for operation and maintenance of the infrastructure, grants to support investment, staff and pension obligations, debt service as well as support granted for the restructuring of the sector. Initial 2001 data have been updated and extended for a period from 1995 to 2003 (see NERA 2005), covering public budget contributions to railways in nine European countries (see Figure 1). Data sources comprise annual reports, budget reports and the International Railway Statistics published by the International Union of Railways (UIC).

Even though the data has been used in several studies and publications (see, e.g. Perkins 2005 or Dehornoy 2011), we find that the database of NERA (2004) has several shortcomings when trying to compare government support between countries. For example, it covers only support to operators and infrastructure managers while other entities or institutions involved in infrastructure or transport provision are often not included. Staff and pension obligations, for instance, are often paid to institutions outside the administrative area of companies and are therefore not included in the dataset. However, in some countries like Switzerland those expenses have to be borne by the companies itself without additional government contributions.

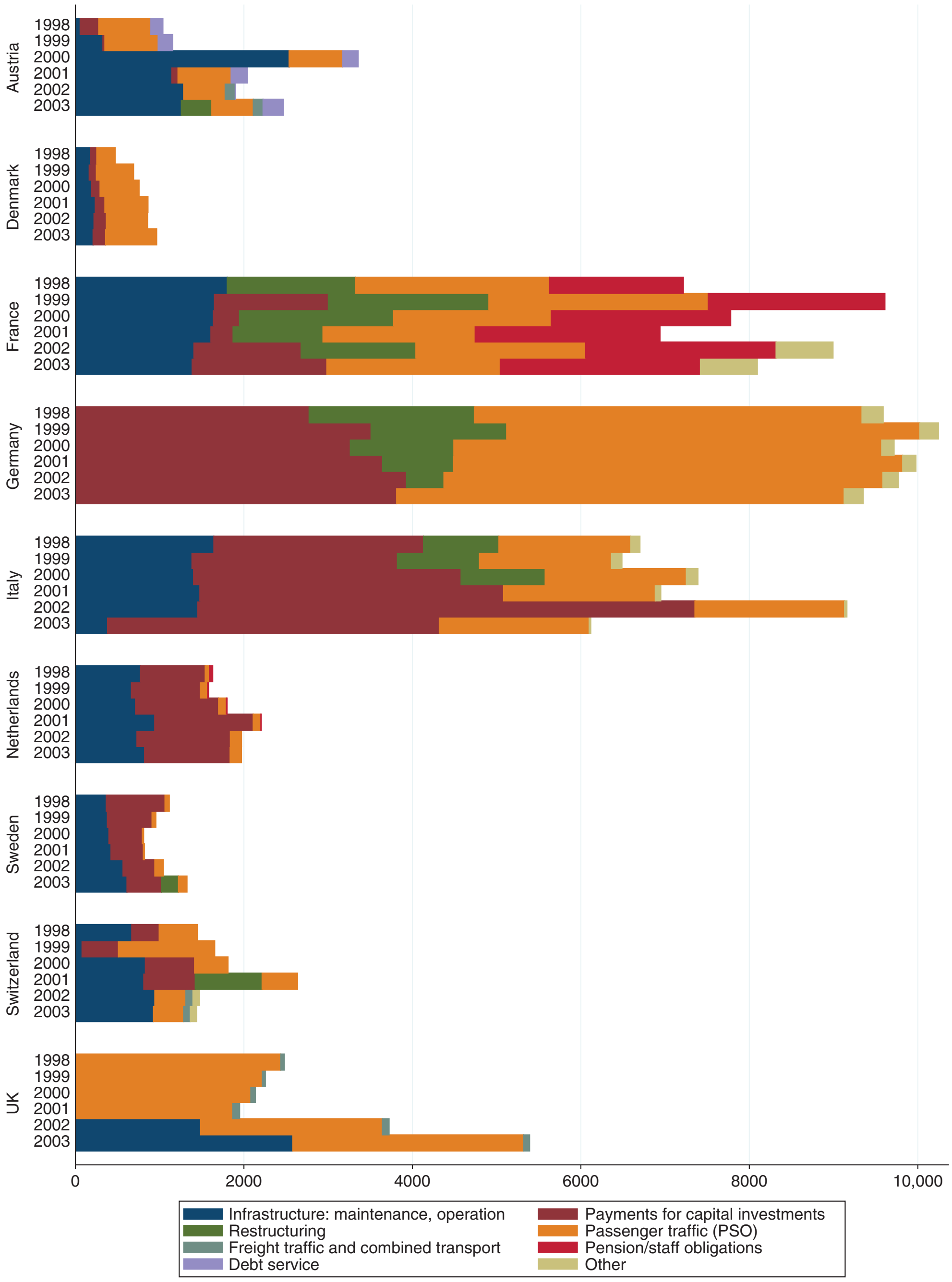


Figure 1: NERA Estimates of Public Budget Contributions 1998–2003.
Source: NERA (2004, 2005). Figures in millions of Euro at current prices.

The exclusion of payments might be more problematic when looking at infrastructure investments. For example, in Switzerland large infrastructure projects are financed from special budget funds outside public accounts. When examining the effectiveness of government support or the performance of a railway system those payments should be included into the analysis, since they might be able to explain differences between countries. Furthermore, the database of NERA (2004) contains only support paid to the main railway companies, mostly incumbents, while payments to other companies are not included. However, since we are most interested in a country comparison it is crucial to consider industry data rather than data for incumbents and main institutions. This is in particular important since the number of private companies taking over public transport contracts tend to be increasing in some countries. These points should therefore be kept in mind when updating data on public contributions.

2.2 Recent Studies and Databases

Apart from the comprehensive database of NERA (2004) and the State aid estimates of the European Commission,[1] data on government support has been collected for a number of analyses who compare data between various countries and for various periods. Typically data is gathered from financial reports of receiving as well as granting institutions in each country. Moreover, sector reports and publications of ministries in charge, regulatory authorities or organizations like the International Union of Railways are usually consulted to obtain a more detailed overview of the financing structure in each country. However, coverage and level of aggregation of data differ between studies as well as between analyzed countries. Data most commonly contain only contributions paid directly to transport operators and infrastructure managers, whereas government loans and support granted via special purpose entities are not always included. Furthermore, it is not always clear to which extent government payments for the restructuring of the sector or other expenditures are accounted for or whether payments to competitors of former state-owned enterprises have been included. We find that a detailed description of contained data, data gaps or the applied collection process is often missing. This hampers the comparison of data and outcomes between studies. The following paragraphs contain a review of selected studies that analyzed government support to the rail sector in European countries.

1 See State aid Scoreboard of the European Commission 2014, Non-crisis aid, Table "Subsidies to the railway sector", available online: https://goo.gl/fKVg5T. Since 2012 data is broken down in Public Service Compensation and pensions as well as infrastructure and other aid. Previous data has been stated in aggregated figures only (see Table "Subsidies to the railway sector", available online: https://goo.gl/NJmNi4).

Nash et al. (2011) compared the overall level of government spending for Great Britain, Sweden and Germany for a period from 1997 to 2007. The authors tried to adjust for different ways of giving support for investment – by grants in Sweden and Germany but by contributing to the cost of servicing debt in Britain. Dehornoy (2011) compared the development of public funding between Germany, UK, France, Sweden and Switzerland for a period ranging from 1980 to 2010 (depending on country). The author tried to break down payments in accordance to the classification of NERA (2004), however, in their final analysis they had to use aggregated figures since detailed comparable data was not available for all countries. Dehornoy (2011) furthermore mention the importance of a consistent comparison of support, in particular of payments for the restructuring and debt relief programs that reduce the cost of some undertakings while other undertakings have not been exempted from these additional burden. We will address these consideration when analyzing debt service payments (see Section 5.2.2).

Deville and Verduyn (2012) conducted a detailed analysis of the financial situation of the railways in Belgium, Germany, France and the Netherlands, though, data is not being used to compare the development of government funding between countries. Steer Davis Gleave (2005) compared the level of state funding of 18 EU countries, Norway and Switzerland for a period from 2000 to 2004. However, a detailed breakdown of support (categories infrastructure development and Public Service Contributions) is only available for the year 2003 and not for all countries. RGL Forensics, Frontier Economics and Aecom (2009) prepared a detailed report of the financial situation of transport undertakings and infrastructure managers to the European Commission covering a period from 2004 to 2007. They provide detailed data on public funding that national railways received from the government, however, companies have been analyzed independently without attempts being made to assess differences in the financing structure between countries. The European Commission (2007: p. 111) published a comparison of the development of financial contributions for infrastructure operation, maintenance, renewals and construction for an period from 1996 to 2006. Unfortunately they do not provide an exact breakdown of support figures.

Less comprehensive analyses contain data for single years or provide only figures on the relative development of support without an exact breakdown of payments. Arrigo and Di Foggia (2013), for example, analyzed schemes and levels of government support to the railway sectors in Germany, UK, Sweden, Italy and France for different years. However, only for Italy payments were broken down into operating support and capital grants (covering a period from 1997 to 2011). Arrigo and Di Foggia (2014) extends and deepens the analysis to cover public spending to railways in France, Germany, Great-Britain and Sweden for a period ranging from 1992 to 2012. However, data is only comparable to a small extent.

In their country comparison the authors therefore considered only aggregated support figures.

In their report to the European Commission ECORYS Nederland BV (2006) compared the financial situation of 23 EU countries, Norway and Switzerland. Information on government support is not available for all countries and the level of aggregation of funding data varies between countries. Nash (2010) analyzed state spending on rail infrastructure between 25 European countries using 2006 data provided by the Community of European Railway and Infrastructure Companies (CER). For their EVES-Rail study van de Velde et al. (2012) compared the development of government support per transport unit for Germany, UK, Switzerland, Netherlands and France over a period from 2000 to 2010. Figures include operating and investment grants, pension obligations and health insurance costs. Dividends that some groups pay back to the government have been subtracted. Nevertheless, no attempts are made to analyze the breakdown of support and contributions are only stated in an aggregated form. However, the EVES-Rail study exposed some problems with the UIC data, not just that it only covers incumbents but that in some cases (including Germany and France) it includes costs of their operations in other countries.

Finger et al. (2015) compared the financial structure of 10 European countries in 2012. Data is taken from the Rail Market Monitoring Scheme (RMMS) of the European Commission (2014), Eurostat, the UIC and other publications. The authors discuss the interaction of public contributions to infrastructure and transport undertakings as well as the influence of different approaches of access charge and fare pricing on the focus of government contributions. Using the data collected in our paper, the analysis of Finger et al. (2015) could be extended for a longer period. This would allow to examine changes in the financing structure of each railway system and their impact on the performance and output of the industry over time.

3 Construction of the Data Set

3.1 Identification of Relevant Payments

The collection of financial data, in particular in regard to public contributions, is an elaborate process. This is due to several reasons. Railways consist of rather complex financing structures. Government support is paid on different stages by the means of different methods and might be complementing or substituting other sources of funds. European legislation does not provide a clear classification of government support/funding instruments and detailed data is not available for the period after the liberalization of the European railway sector.

Furthermore, it is often not easy to identify the "true" extent of government support. For example, if the government is the biggest shareholder of railway companies, it is challenging to determine the effect of government influence on investment decision of the entity and the financing of these (see, e.g. Steer Davis Gleave 2005: pp. 72–73). Especially Public Private Partnerships (PPPs) are said to be used to transfer debt out of public accounts and should be stated as government investments (see Engel et al. 2010: p. 62), however these financing instruments are often not accounted for in existing studies. In practice the impact of so called indirect support is difficult to measure (see, e.g. Valsecchi et al. 2009: p. 17), since no standardized demarcation of indirect support exists. Some researchers suppose that indirect support accounts only for a comparably small portion of total support (see Best et al. 2006: pp. 6–13) while others estimate the level of indirect funding to up to 30 percent of declared State aid (see NERA 2004: p. 118). However, literature focuses mainly on contributions that can be observed directly. We will follow this approach to be in line with existing studies and to avoid estimation errors. Thus, in this study the terms "government support" or "public contribution" refer to any direct government expenditure which is targeted to the field of transport service and infrastructure provision as well as the restructuring of the sector or other fields mentioned in the next section.

Following suggestions of the OECD (2010: pp. 83–86) who published a guide on measuring public support to the agricultural sector, we captured relevant government payments using a three-step identification process. We applied this procedure to overcome shortcomings of previous studies who mainly concentrated on the biggest railway companies and excluded support paid to special purpose entities or investment undertaken directly by the government. It allows to reveal the overall level of government funding rather than the support to single companies. First, all institutions that are involved in public funding of the railway system will be identified. Second, all administrative levels will be covered, since funding might pass several stages before being granted to the final beneficiary. Third, all public finance instruments will be considered, even those that are organized outside the national budgets, for example funding for investment that is provided through special purpose entities.

3.2 Classification of Contributions

To achieve comparability of the collected data between countries we classify the identified payments to one of the categories mentioned below. The categories reflect the main areas of government funding and are derived from existing studies (in particular NERA 2004, 2005), the organizational structure of the sector as well as the legislative background and financing practice as discussed in the previous section. We have slightly modified the classification of NERA (2004: pp. 61–68) to account for additional types of support. Our analysis will focus on the following categories of government contributions: (1) infrastructure revenue contribution (support for management, operation and maintenance of the network) that are paid to complement/substitute revenue from access charges, (2) infrastructure investment contributions for (a) replacement investments as well as (b) for the expansion of the existing network and new construction projects, (3) Public Service Compensation (PSC), (4) investment contributions for the purchase of rolling stock, (5) pension and staff obligations, (6) payments to reduce the indebtedness of undertakings, debt service payments, compensation for interest expenses as well as the takeover of historical debt and (7) obligations related to restructuring. In the following we discuss these definitions in detail.

Infrastructure revenue contributions (1) refer to public funding that covers operating costs of the railway infrastructure. These payments are often classified as operating income in the profit and loss account of the infrastructure managers and are subject to management contracts or compensate the IM for specific operating expenses. Government support for the task of establishing, managing and maintaining the railway infrastructure are subject to Article 8 of Parliament and Council Directive (EU) No 2012/34 (OJ L 343/32) establishing a single European railway area. However, it is not explicitly stated in which forms financing can be granted. According to a statement of the European Commission (2008b) the relationship between each Member State and its infrastructure manager for the funding of maintenance and modernization of the railway infrastructure should be subject to a multi-annual-contract. Government support thereby shall complement revenue from user charges, i.e. should be treated as additional income from operation. However, Council Directive (EU) No 2012/34 also gives Member States the right to demand the infrastructure manager to balance its profit and loss account without government support, i.e. to cover operational expenditures only from user charges.

Council Directive (EU) No 2012/34 furthermore points out that Member States may provide the infrastructure manager with financing, in particular to cover new investments. In our analysis we split contributions for infrastructure investments (2) into two sub-categories, support for replacement investments (a) and support for expansion and new construction (b). This involves grants for the renewal of the existing railway infrastructure (a) as well as grants for investments in new railway infrastructure and expansion of the existing network. This will allow examining to which degree the government supports the development of the infrastructure in terms of focusing either on the preservation or/and on the extension of existing infrastructure. Moreover, the influence of governmental support on different stages of the provision of railway services can be studied to a greater extent. For instance, especially support for infrastructure operation and maintenance as well as support for replacement investments lowers the access charge and therefore the cost of transport undertakings, i.e. necessary Public Service Compensation (see NERA 2004: pp. 64–66), while other countries might support the infrastructure indirectly by higher payments for Public Service Obligations (PSO).

The Category Public Service Compensation (3) refers to contractual payments and revenue contributions that are subject to the public provision of transport services. Public services are most commonly referring to regional transport services, nevertheless, some long-distance or freight connections are also seen as economically desirable and receive government support. However, it should be noted that public service activities are not necessarily non-profitable. Due to their importance and the use of cross-subsidization to reduce the amount of necessary government funds, the decision-making power usually remains with the responsible authority. Public service activities are typically contracted to a commercial operator. Parliament and Council Regulation (EC) 1370/2007 (OJ L 315/1) on public passenger transport services by rail and by road sets the conditions under which authorities can compensate transport operators for costs incurred or can grant exclusive rights in return for the discharge of public service obligations. The Regulation defines that a public service compensation refers to any benefit, particularly financial, that can be granted directly or indirectly to operators (see article 2 g). Since the measurement of non-financial, e.g. indirect, compensations can be only approximate, we focus only on direct compensations as stated in company accounts and budget plans.

According to European legislation support to transport operators outside the scope of a Public Service Obligation may also be granted if it is compatible with the Treaty. Community Guidelines on State aid for railway undertakings (see European Commission 2008a) point out, that financing of rolling stock can contribute to the objective of common interest. Thus, governments may support transport undertakings with funds for the purchase and renewal of rolling stock. Only the costs of the purchase of rolling stock for exclusive use in freight transport are not admissible (see paragraph 34). For all other types of transport strict conditions need to be met to avoid distortions of competition (see paragraphs 31–40). If transport operators receive additional grants for investment in rolling stock, payments are stated in the category rolling stock investment contributions (4). The treatment of rolling stock investment contributions differs to the approach of

NERA (2004), who have not stated them separately but included them into PSC figures.

Payments for retirement as well as redundancies, i.e. government support for staff payments that are not directly related with the operation of infrastructure or transport services, are contained in the category pension obligations (5). Debt service, debt reduction and capital rejections (6) refers to the takeover of historical debt and debt service payments in the form of capital transfers or the compensation of interest expenses. The category obligations related to the restructuring (7) contains government support like the compensation for specific operating cost and compensation for legacy cost that are due to the management of a former state owned enterprise. In the past, support for restructuring was most commonly granted in the form of debt reliefs with the aim of freeing the sector from “historically” grown debt. Government support thereby focuses on relieving operators from additional burden that private operators do not need to bear. Thus, support related not directly to operation and investment is most often justified by promoting equality between former state owned enterprises and other commercial companies [see Council Regulation (EEC) No 1192/69 on common rules for the normalization of the accounts of railway undertakings, 1969, OJ L 156/8]. The degree to which former state owned enterprises have been exempted from historic debt has an influence on the financing cost the undertakings have to bear, and thus influences the amount of necessary funds that have to be recovered either from revenue or from public funds, this is why it is crucial to include these payments into the analysis.

3.3 Sample Selection

Our database contains eight European countries: France, Germany, Great-Britain, Italy, Norway, Spain, Sweden and Switzerland. These countries have been selected for several reasons. According to Nash (2008) the organizational structure of railways can be categorized into three groups: Countries where infrastructure is separated from operations (Great-Britain, Norway, Spain and Sweden), countries that operate their incumbents railways in the form of a holding company with organizational separation (Germany, Italy and Switzerland) as well as countries that have separated infrastructure and operations but subcontracted much of its activity to the incumbent like it is the case in France.[2] The dataset covers all three organizational models and will allow analyzing whether the structure of government contributions depends on the degree of separation.

2 It should be noted that France moved back to an holding structure in the beginning of 2015.

Table 1: Country Characteristics (Average Values 2001–2016).

	Area km²	Population		Billion		Network		Modal share	
		mio.	Density	Tkm	Pkm	km	Density	Pass.	Freight
CH	41,285	7.18	192	11.54	17.37	3483	84	16.8	44.8
DE	357,125	76.63	229	102.7	81.36	35,303	99	7.9	21.4
ES	505,988	41.85	89	10.62	22.47	15,381	30	5.5	4.8
FR	632,881	59.84	101	39.35	81.68	29,448	47	9.3	16.2
IT	301,483	54.98	198	21.16	49.11	16,744	56	5.9	11.6
NO	323,782	4.48	16	3.279	2.965	4188	13	4.6	14.7
SE	440,624	8.66	23	21.39	10.49	11,042	25	8.1	36.3
UK	248,528	57.82	255	20.77	52.02	16,433	66	6.9	11.7

Population density is expressed in inhabitants per km² while network density is expressed in km per 1000 km².

Also, it was important to select countries that have comparable structures, either in terms of country or sector size. For example, Norway and Switzerland have a comparable network length (see Table 1 for select data of the respective countries). The same holds for Germany and France. Sweden and Switzerland have a comparably high modal share of freight while Germany and Great-Britain have a comparable population density. The smallest country in our sample is Switzerland with an area of 41,285 km². However, in terms of traffic units it can compete with the Spanish railway sector. Norway has the smallest population density (16 inhabitants/km²) as well as the lowest modal share of passenger transport (4.6 percent) in the sample. Factors like population density and network length might influence the cost and thus the financial needs for providing services and should therefore be included in future empirical analyses.

3.4 Data Sources and Quality

The quality and the coverage of available data on government contributions has increased in recent years. To maintain comparability of data over the whole observation period minor adjustments had to be made. Nevertheless, there are still large gaps in data for some countries. Furthermore, differences in data availability, scope and consistency persist between countries that have been considered in our study. There are also large differences concerning the accounting treatment of government support. It is clear that due to the complexity of the sector a full coverage of data cannot be achieved. However, this problem is exacerbated by the fact that no major attempts have been made at the political level to standardize the reporting of government funds being spent on the sector, even though they account for a significant portion of Government budgets in each country. Especially in times of tightening budgets the debate on the future design of the railway sector in Europe might benefit from an in depth analysis of differences in the funding structure and return on government funds.

For France we compiled data mainly from the annual transport statistic published by the French Ministry of Transport. Although data of the Ministry of Transport differ slightly from the payments stated in annual reports of SNCF and RFF they are still advantageous, since data is available for a longer period and allows for a detailed breakdown of support figures. The amount of compensation for infrastructure operation in our dataset is comparable to the figures in NERA (2004, 2005). Nevertheless, we obtained different estimates for the height of the granted infrastructure investment contribution, Public Service Compensation and pension obligations. Dehornoy (2011) and Deville and Verduyn (2012), who analyzed the development of public contributions to the French railway sector in detail, also used data published by the French Ministry of Transport. Since both publications do not provide detailed information on the exact height of payments included into each of their categories, figures are difficult to compare with our findings. However, the relative development of total support estimates is similar.

Data for the German railway sector has been mainly taken from the budget plans of the Ministry of Transport. However, some adjustments had to be made. We supplemented data by comparing data with figures published by NERA (2004, 2005), the financial reports of Deutsche Bahn, the outcomes of the studies of RGL Forensics, Frontier Economics and Aecom (2009), Dehornoy (2011), Nash et al. (2011) and Deville and Verduyn (2012) as well as with information and data obtained from correspondence with the Federal Network Agency (Bundesnetzagentur) and Deutsche Bahn. In addition, two publicly available statements of the Federal Government concerning the use of transfers to the Federal States have been consulted to trace data gaps and to estimate the approximate amount of support for public services.

Data for Great Britain has been taken from a data-base on Government support to the British rail industry[3] that is updated annually by the Department for Transport, Transport Scotland and the Welsh Government. We complemented data with information from the National Rail Trend Yearbooks which are published by the Office of Rail Regulation, data on payments to TOCs collected by the Department for Transport and the Office of Rail Regulation, the annual reports and accounts of the Department for Transport, the business plans of the Department for Transport, a report of the National Audit Office on the Department for

3 See Table "Government support to the rail industry: annual from 1985/1986", available online: https://goo.gl/2dFWDx.

Transport, the Central Government Supply Estimates, GB rail industry financial information published by the Office of Rail Regulation as well as the annual reports and regulatory financial statements of Network Rail. It should be noted that data for Great-Britain refer to financial years only. Data of all other countries refer to calendar years. Furthermore, figures are net of received performance payments, whereas figures for other countries contain gross data.

Data for Italy has been mainly taken from the financial statements of FS Italiane and its subsidiaries. Trenitalia SpA provided us with data on Public Service Compensation for regional and freight transport services as well as contributions for the purchase of rolling stock. Support to other companies and regional operators is missing in our data set. We are not sure about the approximate level of support granted to other entities since no detailed information was available to us. We used publications of Arrigo and Di Foggia (2013, 2014), NERA (2004, 2005) and RGL Forensics, Frontier Economics and Aecom (2009) to compare our estimates for each category. Data for some categories of our framework were not available to us for all years, therefore we had to exclude government support for 2001–2003 from our final figures.

The Norwegian Department of Public and Rail Transport of the Ministry of Transport and Communications provided us with detailed data on support to NSB as well as to Jernbaneverket. We complemented the data with figures from the annual reports of the NSB Group and NSB AS as well as with data from the annual reports of Jernbaneverket. Comparable support figures were only available in the publication of NERA (2004) for the year 2001. The height of support in each category as well as our total support estimates are in line with our findings. Slight differences could be explained by currency conversion from NOK into Euro using different exchange rates.

Data for the Spanish railway sector has been taken from budget plans of the State as well as from annual reports of RENFE, Renfe-Operadora and Adif. We obtained additional information from the publication of NERA (2004) as well as from the report of RGL Forensics, Frontier Economics and Aecom (2009). In Spain public support for investments is transferred in the form of equity contributions to both, Renfe-Operadora and ADIF/GIF. However, these payments are officially not stated as investment grants in the government expenditure report or the annual reports of the companies, but are used to finance a large part of investments. Eurostat (2008) has criticized the methodological treatment of these payments (see also Instituto Nacional de Estadistica 2007), we have therefore included the payments in our investment contribution figures. Due to the restructuring of the sector in the end of 2004, data availability and consistency problems persist. Coverage of our data differs between the period before and after the restructuring.

Our data set for Sweden is based on support figures submitted by Trafikverket, the annual reports of Banverket, Trafikverket, Green Cargo as well as SJ. We furthermore gathered information from publications of Trafikanalys, a Swedish government agency who is responsible for the publication of official statistic in the transport sector, the budget plans of the Swedish government as well as the Sector Report of Banverket (2008). However, we have not been able to complete data for the whole observation period. Furthermore, detailed data on Public Service Obligations for regional rail transport was not available. Trafikanalys publishes an annual statistic on local and regional public transport (Stockholm region missing); however, payments for the provision of services are not broken down by mode of transport. PSC for local and regional rail services have therefore been calculated by subtracting total revenues from total costs of rail operators. There seems to be a lack of data availability since other authors had similar problems to estimate PSC (see, e.g. Dehornoy 2011). Data on procured interregional rail services is available in the accounts of Trafikverket.

Support figures for the Swiss railway sector have been found in the annual reports/financial statements of the SBB Group, data published and submitted by the Swiss Ministry of Transport, data from the Swiss Statistical Office as well as publications of the Swiss government. Even though a large number of statistical databases exist for the Swiss railway sector, the completion of our data set was not possible without accepting some data gaps. One problem was to determine the amount of Public Service Compensation since the responsibility for public services is shared between the State and the regional authorities. Furthermore, there is a break in time series data between 2006 and 2007 due to changes in the support scheme. It should be noted that the Swiss Statistical Office publishes a detailed account for the railway sector (see Bundesamt für Statistik BFS 2013). Unfortunately, data is too aggregated to use it for our database. Concerning the coverage of our data we believe that the figures that we have collected for the Swiss rail sector undervalue the actual amount of government compensation to a certain extent. Therefore, all estimates, for the swiss as well as for all other countries, should be treated with care.

4 Public Budget Contributions

4.1 Overview

In order to be able to compare the absolute level of public budget contribution between countries we have used GDP Purchasing Power Parities (PPP) to normalize data. This approach is consistent with the studies of NERA (2005), Dehornoy

(2011) and van de Velde et al. (2012) who also used PPP adjusted estimates of public budget contributions.

We find that there are large differences both concerning the level of government support and concerning the structure of government contributions (Figure 2). The development of payments also differs to a great extent between

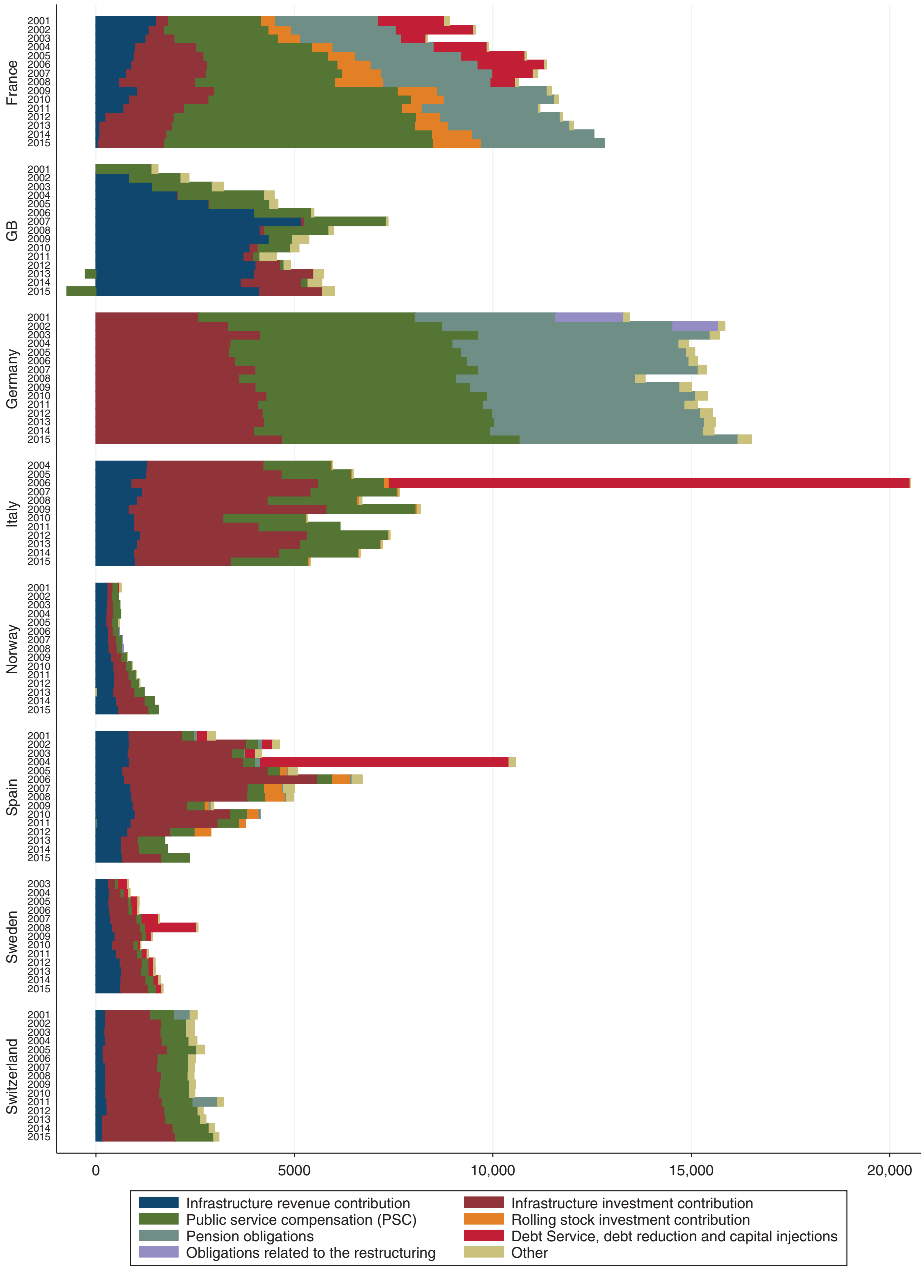


Figure 2: Development and Breakdown of Public Budget Contributions 2001–2015. Figures in millions of PPP adjusted Euros.

countries. Taking into account all categories of support, the German railway sector receives the highest contributions in absolute terms, followed by the French sector. The Italian and British as well as the Norwegian and the Swedish railway sector receive contributions on a comparable level, Spain as well as Switzerland are in between the four.

While support figures evolve relatively stable in France, Germany, Norway, Sweden and Switzerland, there is a lot of volatility in the data for Great-Britain, Italy and Spain. In Great-Britain we observe a strong increase in (infrastructure revenue) contributions since the foundation of Network Rail in 2002. Figures reach their peak in 2007, since 2012 contributions start to climb again. The development is being dampened by repayments of franchises. For 2013 and 2015, there are even surpluses from the franchise contracts.

In Italy contributions reached their peak in 2009, in Spain in 2006. There have been large transfers for the takeover of debt in 2006 (Italy) and 2004 (Spain). In 2008 the Swedish sector received additional funds for the repayment of loans. Till 2008 the French railway sector has also been receiving equity contributions on a regular basis. Onetime payments also occurred in Germany, where there have been transfers for the restructuring of the sector in 2001 and 2002, as well as in Switzerland in 2001 and 2010 in the form of transfers to the pension fund. The French and the German railway sector receive payments to offset the additional burden which is caused by former having been run as state enterprises. This applies in particular to pension payments.

All main infrastructure managers except the German national infrastructure operator, receive contributions for operation which are usually treated as revenue in the income statement. In France payments have been decreasing since 2001, at the same time the compensation for public services has been increasing. In all countries the railway sector has been receiving investment contributions for enhancement of the infrastructure. Contributions for rolling stock investments have been only granted in France, Spain and Italy. However, in some countries additional funds are provided from the budget for public services.

4.2 Relative Development

To adjust for the country and sector size we will use passenger-ton-kilometer (sum of passenger and ton-kilometer) as well as the number of inhabitants in each country as indicators. Passenger-ton-kilometers are one of the most commonly used output indicator that shows the amount of traffic units produced in each country. It should be noted that there are other factors that might explain differences between the level of support, i.e. the economic geography of the country. In a country with dense inter city flows passengers or long distance flows of freight, many services will be profitable and help pay for the infrastructure; in a less densely populated country it may be that all services need support and cannot contribute to the cost of the infrastructure. Also, passenger-ton-kilometer may distort results as freight is more likely to be profitable than passenger traffic, so countries with high levels of freight traffic will appear to have lower support per unit of traffic. Therefore we use the number of inhabitants as a second indicator that allows to adjust data for differences in the size of the country. Furthermore, it is an important measure of taxpayers cost of running the national railways.

To increase the comparability of support figures we exclude all payments that are not directly related to the operation of transport services and the infrastructure or that are not used to finance investments. In the following sections public budget contributions therefore refer only to (1) infrastructure revenue contributions (support for management, operation and maintenance of the network), (2) infrastructure investment contributions for replacement investments as well as for the expansion of the existing network and new construction projects, (3) Public Service Compensations (PSC) as well as investment contributions for the purchase of rolling stock (4). At this point of the analysis, we exclude items (5) pension and staff obligations, (6) debt service and capital injections and (7) contributions related to the restructuring from the analysis. However,it should be noted that especially the way infrastructure investments are financed, might influence cost and thus the necessary level of support. In Section 5.2 we will therefore discuss the role of financing of investments from debt issued by the infrastructure manager as well as debt service payments and capital injections in more detail.

Support per passenger-ton-kilometer (ptkm) is in a range between 0.02 and 0.22 Euro over the observation period. Thus, on average each output unit is supported with 0.08 Euro. Sweden grants the lowest contribution per ptkm while Norway has become the country paying the highest contribution per passenger-ton-kilometer. The average annual contribution per passenger-ton-kilometer varies between 0.03 Euro/ptkm (Sweden) and 0.14 Euro/ptkm (Norway). Some countries have undergone significant changes in the level of contribution per passenger-ton-kilometer during the years. Support to the Spanish railway sector has decreased from more than 0.18 Euro/ptkm in 2006 to 0.06 Euro/ptkm in 2015 while support to the Norwegian railway sector increased from 0.10 Euro/ptkm in 2001 to 0.22 Euro/ptkm in 2015. Support per

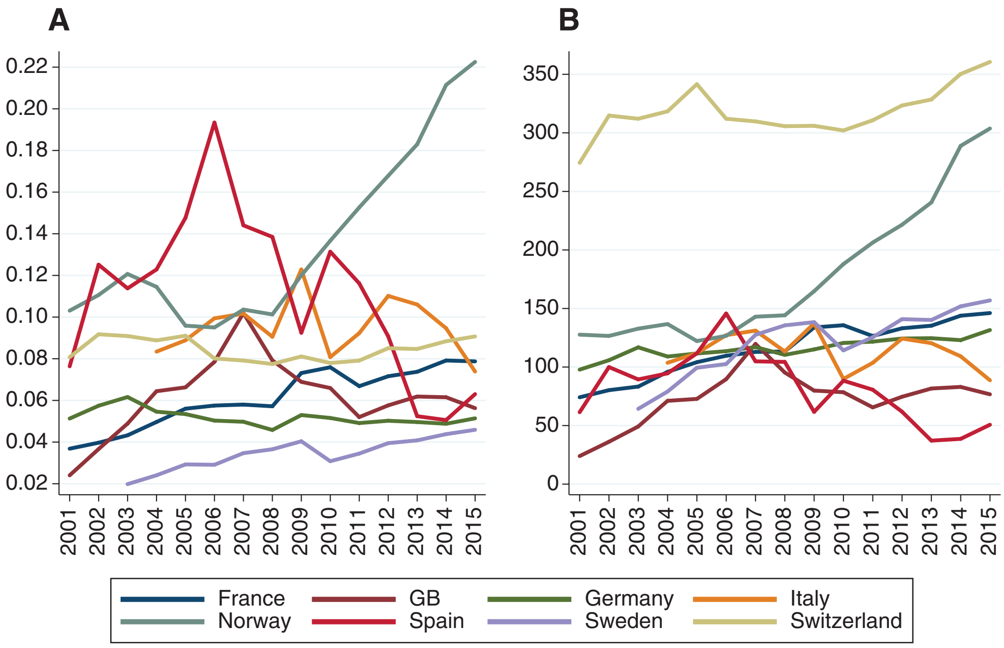


Figure 3: Government Support for Infrastructure Investments, Operation, PSC and Investments in Rolling Stock 2001–2015.
(A) €PPP/ptkm, (B) €PPP/inhab.

passenger-ton-kilometer to the German and Swiss railway sector stayed relatively constant. Payments to the British railway sector have reached their peak in 2007 and tend to decrease since that point. Payments to the Italian railway sector are also characterized by a high volatility, tending to decrease during the last periods. Support to the French and Swedish railway sector has been increasing over the observation period (see Figure 3).

Contributions per inhabitant are in a range between 24 and 361 Euro over the whole observation period. Great-Britain has granted the smallest annual average contribution to its national railway sector (73 Euro/inhabitant) while the Swiss railway sector has received the highest average contribution over the whole observation period (318 Euro/inhabitant). Like the level of support per passenger-ton-kilometer, in some countries taxpayer cost vary to a great extent over the observation period, while in some countries support figures evolve relatively constant over time. Support payments to the Swiss railway sector increased by almost 100 Euro/inhabitant. Support to the Norwegian sector has doubled. Payments in France, Germany and Sweden have been increasing while support to the British, Italian and Swedish sector tend to decrease in most recent years.

5 Applied Financing Models

5.1 Financing of Operation

5.1.1 Ratio Analysis

Similar to Finger et al. (2015) we identified a tendency to mainly two financing models. In our data-set there are countries that either focus their support payments on the operation of the infrastructure, which implies lower network charges and thus a lower amount of necessary Public Service Compensation (PSC); or they focus support payments on transport services with a higher degree of cost coverage of access charges and thus lower operating contributions that need to be paid to the infrastructure manager. However, we also find that in some countries no clear tendency can be observed. These countries apply a "hybrid"-like financing approach where the infrastructure manager as well as transport operators receive an almost equal share of operating contributions. It should be noted that higher track access charges do not simply reduce the size of payments for Public Service Obligations (PSO); they also reduce scope for commercial services. We will discuss differences in the focus of government support as well as changes during the observation period as well its implications on cost coverage and financial sustainability in more detail in the following sections.

Figure 4 shows the amount of granted Public Service Compensation per passenger-km in relation to the amount of infrastructure revenue contribution granted per passenger-ton-kilometer in 2014. One can observe a negative relationship between Public Service Compensations and infrastructure revenue contribution. Countries with a comparably small level of contributions for Public Service Obligations support the operation of the network to a greater extent while countries that are granting a high compensation for passenger transport services support the operation of the network to a smaller extent.

By comparing public budget contributions for the operation of the infrastructure as well as Public Service Compensations relative to the traffic volume we were able to identify three different support schemes. In 2014 Germany, France and Great-Britain are the countries that focus either completely on support for transport services or on support to the infrastructure manager. All other countries use a combination of support for public services and infrastructure operation. As further emphasized in the following chapters, revenue contributions in Great-Britain are also used to cover interest expenses. Considering only the part that is used to cover operating expenses of the infrastructure, Britain should move in direction of the origin in Figure 4.

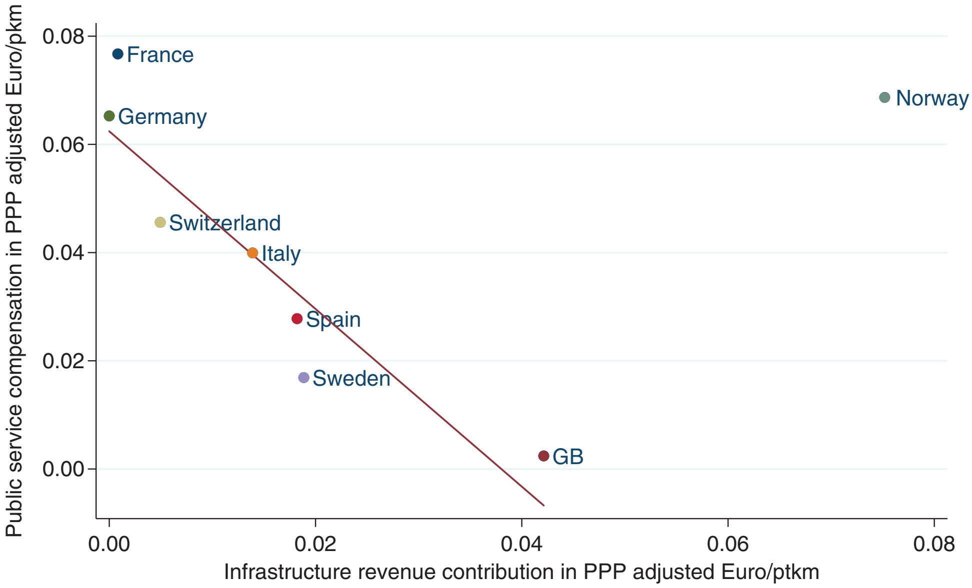


Figure 4: Ratio of PSC and Infrastructure Revenue Contribution in 2014.
Investment contributions and other support categories have been excluded. We furthermore excluded Norway from the calculation of the average ratio, since data is out of range of the other countries.

Support to the Norwegian railway sector lies outside the range of all other countries and has therefore been excluded from the analysis. However, the outlier might be explained by the comparably small amount of passenger-ton-kilometer demanded in relation to the network size and thus higher cost per passenger and ton-kilometer of the Norwegian sector. The Swiss railway sector, that has a similar network size, provides an almost four times higher amount of traffic units compared to Norway. However, one should also keep in mind differences in the population size and density that could explain differences in transport demand. Also, the degree to which access charges cover operating expenses of the infrastructure in Norway is among the lowest in the sample (see Section 5.1.3).

5.1.2 Development of PSC and Revenue Contributions

There are large differences between the levels and development of public budget contributions within the different support categories. Figure 5 shows the granted infrastructure revenue contribution for a period from 2001 to 2015 in PPP adjusted Euro relative to passenger-ton-kilometer and to the number of inhabitants. Figure 6 shows the respective development of Public Service Compensation.

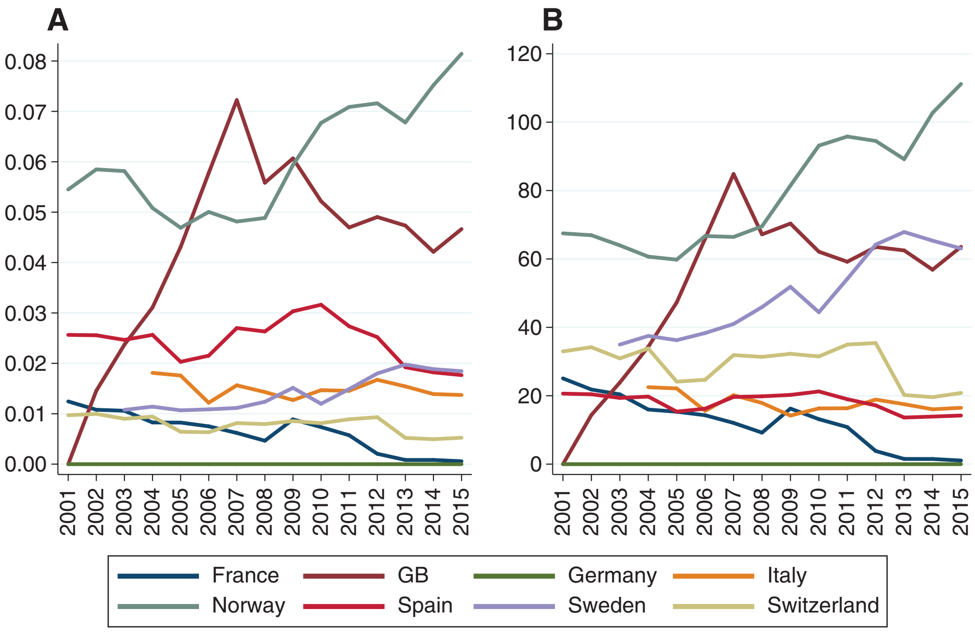


Figure 5: Infrastructure Revenue Contribution 2001–2015.
(A) €PPP/ptkm, (B) €PPP/inhab.

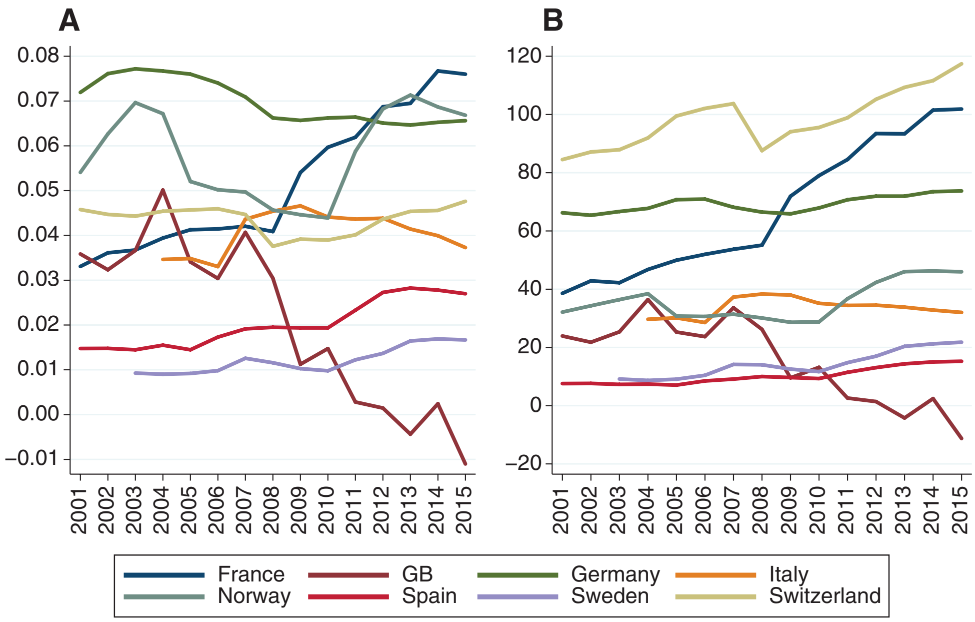


Figure 6: Public Service Compensation 2001–2015.
(A) €PPP/pkm, (B) €PPP/inhab.

The infrastructure revenue contribution which refers to public contributions granted for the operation of the network, varies to a great extent between countries. The British railway sector receives by far the largest contribution in absolute terms, reaching 5600 million PPP adjusted Euro in 2007. In relative terms only the Norwegian railway sector is receiving higher contributions, reaching 0.08 Euro per passenger-ton-kilometer or 111 Euro per inhabitant in 2015. Contributions to the railway sector in Great-Britain have experienced a strong increase in both, absolute and relative terms since the foundation of Network Rail in 2001. After reaching its peak in 2007 contributions tend to decrease in the following years. The average revenue contribution over the whole observation period and over all countries amounted to 0.02 Euro/ptkm. Thus, the infrastructure managers receive an average compensation of approximately 2 cent for each passenger-ton-kilometer that is demanded on their network. Tax payer cost of operating the national network amount to an average of 32.3 Euro per inhabitant. Only the German railway sector finances the operation of the infrastructure entirely from network access charges and therefore receives no revenue contribution.

The average Public Service Compensation amounts to 0.04 PPP adjusted Euro per passenger-kilometer. However, values and development of support differ to a great extent between countries. While the compensation is decreasing over the observation period in Germany and GB it tends to increase in all other countries. Germany is paying the highest average compensation amounting to 7 cents per demanded passenger-kilometer while the Swedish operators receive the smallest amount of compensation per passenger-kilometer amounting to an average of 1 cent over the observation period. The development of Public Service Compensation in France and Great-Britain are exceptional. While the compensation per passenger-kilometer to the French transport sector has more than doubled, the compensation to the British operators has reached values close to and below zero. However, it should be noted that support estimates for Great-Britain are net figures (support net of revenue from franchise allowances), i.e. they might be undervalued compared to the development in other countries. This is particularly true for most recent years because of an increasing amount of performance receipts from franchise holders that offset the paid gross compensation.

5.1.3 Cost Coverage of Access Charges

Operating cost of infrastructure managers (net of interest expenses) are covered to different degrees from access charges, other revenue and government support.

The same holds for cost of operating Public Services where cost are covered from fare revenue and Public Service Compensations. From our ratio analysis in the previous section one would expect, that countries with a focus on infrastructure financing would reach lower cost coverage ratios from access charges since revenue is substituted by government contributions and vice versa. Laurino et al. (2015: p. 209) also assume that the pricing principles vary according to the applied financing model and other country specific characteristics. However, we find that there are also differences between the cost coverage ratios of countries that apply an almost equal financing scheme.

Figure 7 shows the degree to which access charges for the use of the national railway network are covering operating expenses. Underlying data include staff expenses, material purchases, depreciation as well as other expenditures. Countries using a financing scheme with focus on the support of public transport services like Germany and France reach cost coverage ratios of more than 80 percent, while countries which are applying a hybrid-like financing model reach cost coverage ratios of only 20–40 percent. Even though government contributions in Great-Britain are mainly paid in terms of revenue contributions to Network Rail and public services receive no (net) support, the cost coverage ratio of access charges is still higher than in countries with a hybrid-like financing model, amounting to more than 40 percent. This is due to the fact, that a large part of revenue contributions is used to finance the debt of Network Rail

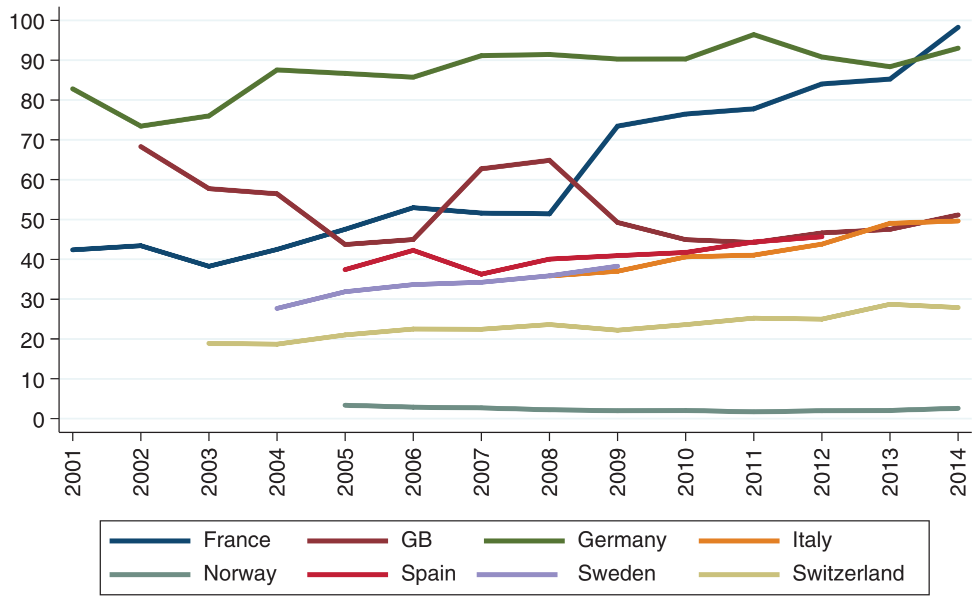


Figure 7: Share of Infrastructure Operating Cost Covered by Access Charges 2001–2014 in Percent.

rather than to finance the operation of the network. The remaining part has to be covered from access charges. When looking at differences in the level of contributions for the operation of the network, one should therefore adjust payments by the part that is used to finance debt or investments.

Revenue from rail access charges in Norway covers only 2–3 percent of operating cost, with revenue contributions being the main source of income. In Italy, Switzerland, Sweden and Spain the degree of cost coverage has been slightly increasing over the observation period, while the cost coverage ratio in France has seen a sharp increase. The French network operator RFF has more than doubled the cost coverage of infrastructure access charges from around 40 percent in 2001 to more than 90 percent in 2014. As mentioned earlier this development is due to a change in the support scheme. However, since the revenue contribution to RFF has been substituted by higher payments to operators of public transport services, total support figures have only slightly changed.

5.2 Financing of Infrastructure Investment

5.2.1 Investment Contributions

Concerning the financing of infrastructure investments two approaches exist. Either investments are financed from government contributions or from funds of the infrastructure manager. The latter typically involves issuing debt. Some infrastructure managers are compensated for the additional burden of financing debt while other IM need to cover these expenses from network charges and revenue contributions.

The governments in the countries that are subject to our study grant funds of different size for infrastructure investments to their national railway sectors. In absolute terms Germany has received the highest average annual contribution over the observation period amounting to 3840 million PPP adjusted Euro. While contributions develop relatively constant in most of countries, there is a lot of volatility in contributions to the Italian and Spanish railway sector (see Figure 8). Tax payers cost for investment in infrastructure amount on average to 62.5 Euro per inhabitant over the whole observation period. The Swiss railway sector receives the highest annual average contribution amounting to 190 Euro per inhabitant, while the British government supports investment in the infrastructure only with an average annual contribution of 6 PPP adjusted Euro per inhabitant. Except Switzerland all other countries grant an annual average contribution that is <65 Euro per inhabitant. However, high contributions in Switzerland might be the

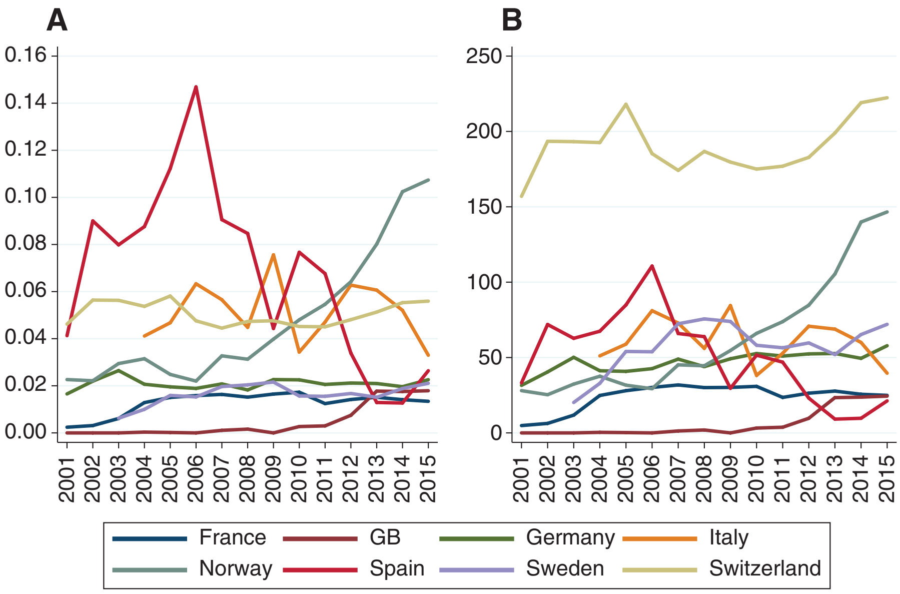


Figure 8: Infrastructure Investment Contribution 2001–2015.
(A) €PPP/ptkm, (B) €PPP/inhab.

result of expensive construction cost since a large number of bridges and tunnels are necessary for the operation of alpine crossing railway lines.

The main funding source for investments in the German infrastructure is the government budget. The Norwegian sector also finances nearly all investments from direct government contributions, while the degree in other countries is between 20 and 60 percent. The share of infrastructure investment that has been financed from direct government contributions has been increasing in France, Germany, Great-Britain and Italy while the degree of government contributions tends to fall for other countries. Especially Spain has reduced the degree of government contribution to a strong extent, from more than 60 percent in 2006 to around 25 percent in 2011. In Great-Britain nearly all investments are financed from other sources than government funds, i.e. funds raised on the capital market.

These differences in the degree of government funding are observable in the amount of debt issued by infrastructure managers. Figure 9 shows the development of non-current liabilities of the main infrastructure managers over a period from 2001 to 2015 in million PPP adjusted Euro. While the liabilities of DB Netze Track (Germany), SBB Infrastructure (Switzerland), RFI (Italy) and Jernbaneverket (Norway) evolve relatively constant over the last 14 years, there is a strong increase in non-current liabilities of the main infrastructure manager of France, Great-Britain and Spain observable. Over the last years debt of infrastructure managers in France, Great-Britain and Spain has increased dramatically, which

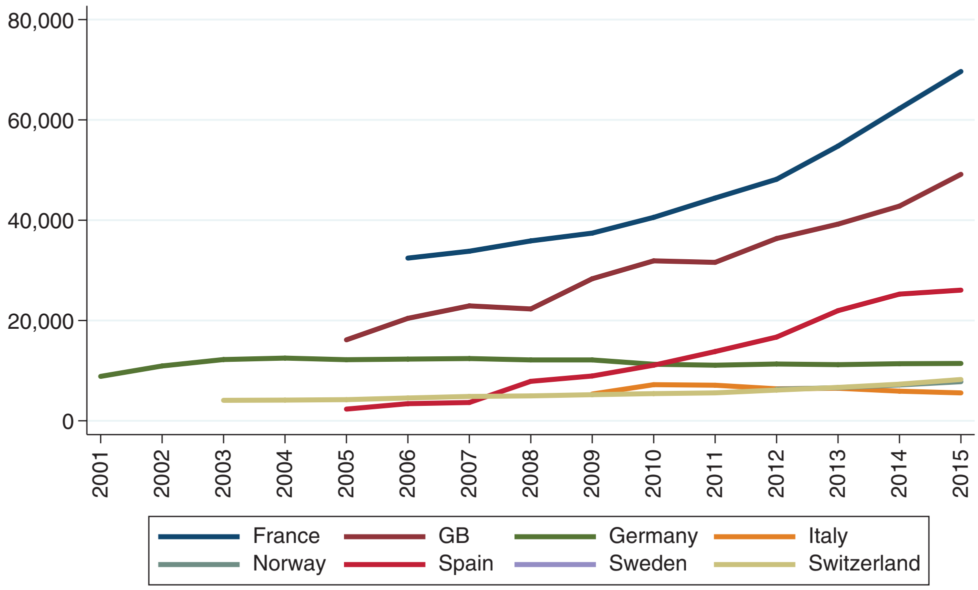


Figure 9: Non-Current Liabilities of Main Infrastructure Managers 2001–2015 in Million PPP Adjusted Euro.
Note: We excluded data for some countries in order to avoid breaks in time series data. Figures for DB Netze Track involve total liabilities (current and non-current). No data was available for Sweden.

in the long-run could limit the financial flexibility of companies and require additional government support or restructuring incentives. NERA (2004: pp. 117–118) even argues that a rising debt level is an indicator for inadequate financing of capital expenditures and that the main goal of government policy should be to ensure sustainable debt levels which they have specified as the degree of indebtedness of other commercial sectors.

5.2.2 Debt Service and Capital Injections

In particular the French and the Spanish rail infrastructure managers have repeatedly received additional government funds to finance their indebtedness (see Figure 10). The Spanish railway group Renfe has been relieved from a large part of its debt before the infrastructure was separated and taken over by ADIF in 2004. RFF, the national infrastructure manager in France, has received grants to reduce its indebtedness between 2004 and 2008 as well as capital injections in 2001 and 2002. However, in recent years no additional funds have been transferred to RFF.

For the Italian High-Capacity project (a railway link between Turin, Milan, Rome and Naples) a joint-stock company called Infrastrutture SpA (ISPA) was

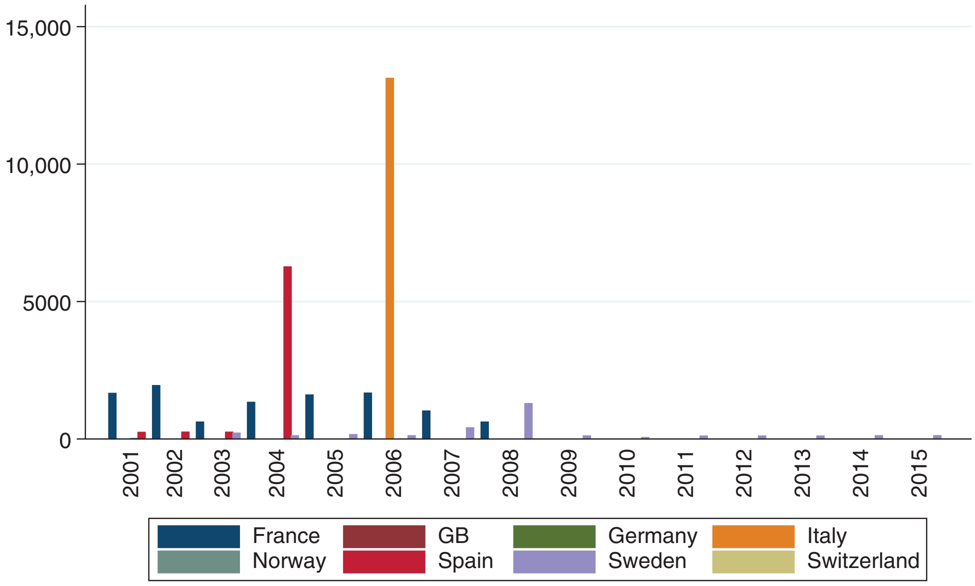


Figure 10: Debt Service, Debt Reduction and Capital Injections 2001–2015 in Million PPP Adjusted Euro.

founded in 2002. ISPA raised funds on the capital market and provided them to TAV, a RFI subsidiary that was responsible for the development of the Italian high-speed network. After the completion of the project the State should have provided RFI-TAV with the amount necessary to cover the difference between the revenues arising from the operation of the new railway link and the debt service owed to ISPA. Since the State was bearing all the risk, it has been decided that the debt issued by ISPA has to be treated as government debt because RFI-TAV will not be able to repay most of the provided loans without additional government support (see Eurostat 2005). The inclusion into public accounts was completed by a debt assumption, i.e. a capital transfers from the Italian government to RFI-TAV in 2006.

Further examples can be taken from Perkins (2005: pp. 8–9), who provides an overview of actions that have been taken to reduce the indebtedness of railway infrastructure and transport undertakings. He shows that in most countries railway sector debt has been taken over by the governments while the debt of the French railway sector remained with RFF. Regarding the financial situation of infrastructure managers there are therefore large differences between the countries. While Sweden and Norway operate and develop their national rail infrastructure by government authorities, other infrastructure managers are run as private companies with the governments being the main shareholder.

Infrastructure managers that carry a large amount of debt must bear higher interest burdens which influences total operating cost and thus the amount of necessary funds which must be covered by access charges or government contributions. Dehornoy (2011: p. 15) tried to incorporate the additional financial burden that is caused by the debt into his support estimates by estimating financing cost. He finds that the debt of the British and French railway sector generate financing costs of around 1.5 billion Euro each while financing costs of the German infrastructure operator DB Track amount to 0.6 billion Euro. On the other hand, an institution like the Federal Railway Fund (BEV) in Germany relieves some undertakings from additional cost that other undertakings need to bear by themselves. In the German case the annual benefit is estimated to an equivalent of 0.2 billion Euro. However, his results are preliminary and need to be treated with care.

To sum up, the way in which support is given may affect the overall level of costs and thus the necessary level of support to achieve particular objectives. Supporting debt servicing rather than providing grants for infrastructure or rolling stock investment affects the overall level of costs; the latter makes the level of current costs seem higher. For instance, debt servicing is a large part of the costs of Network Rail. Similarly, past write offs of debts or takeover of pension obligations will make current costs and support seem lower. These considerations must be taken into account when interpreting the results from the previous sections.

5.3 Summary

There are large differences between the financing of infrastructure operation, public services and investment. Table 2 summarizes our findings about the financing structures of the railways in France, Germany, Italy, Norway, Spain, Sweden, Switzerland and Great-Britain for 2013. The main characteristics of each country will be described below.

With the exception of Germany all countries in our data-set grant revenue contributions to their national infrastructure manager. In Germany access charges are the main source of funding for infrastructure operation (cost coverage 90 percent in 2014), while investments are mainly financed from direct government contributions as well as funds of DB Netze Track. France finances the operation of the infrastructure from a combination of government funds as well as access charges. Infrastructure investment are financed from government contributions and through debt issued by RFF. The additional financial burden is covered from revenue and government contributions. However, between 2001

Table 2: Financing Structure of Analysed Railway Systems in 2014 Grouped by the Degree of Separation.

	Separated				Holding			Hybrid
	GB	ES	SE	NO	CH	DE	IT	FR
Infrastructure								
Revenue contribution	X	X	X	X	X		X	X
Investment contribution	(X)	X	X	X	X	X	X	X
Financing of investments from debt issued by the IM	X	X	X					X
Transport								
PSC	(X)	X	X	X	X	X	X	X
Investment contribution rolling stock		(X)				(X)	(X)	X
Other								
Pension obligation						X		X
Debt service, debt reduction and capital injections	(X)		X					
Focus of operating grants								
Infrastructure	X							
Public services						X		X
Hybrid		X	X	X	X		X	

(X) refers to indirect means of government support. For example, investment contributions for the purchase of rolling stock are often contained in public service contributions and are not stated seperately in our database. Furthermore, public services in GB are profitable by means that the performance receipts from franchise holders exceed gross payments.

and 2007 RFF has made large losses which had an influence on the financial sustainability of the company.

Italy, Spain, Sweden and Switzerland use a hybrid model where both, transport service operators and infrastructure managers, receive an almost equal share of total financial contributions. Investments in the Spanish railway infrastructure are financed using a combination of debt issued by Adif and government contributions. Investment expenditures of the Swedish Trafikverket are financed from government contributions as well as (state) loans. Financial expenses as well as the repayment of loans are partly financed from additional contributions. In Switzerland large infrastructure investments are financed from funds of a special purpose entity. Other investments are financed from direct contributions, (State) loans as well as funds of SBB Infrastructure.

The British sector is the only one which uses a model where the majority of government funding is granted to the infrastructure. Infrastructure investments are almost entirely financed through debt issued by Network Rail. Related debt service payments are financed from revenue contributions. Therefore Great-Britain is the country receiving the highest infrastructure revenue contribution in our data-set. Nevertheless, cost coverage of access charges reached around 50 percent in 2014. In Norway infrastructure operation is almost entirely financed from government contributions. Even though investments in infrastructure are mainly financed from direct government contributions, the indebtedness of Jernbaneverket has been slightly increasing in recent years. Compensation payments for public services in Norway are comparably high.

We are not able to observe a clear connection between the degree of separation and the choice of a specific financing model. However, four of five countries that have completely separated infrastructure and operations or use a hybrid-like approach are financing infrastructure investments from debt issued by the infrastructure manager (Great-Britain, Spain, Sweden, France). In Great-Britain, Spain and France liabilities of national infrastructure managers continue growing. Whereas, in countries where both, infrastructure and operations are owned by a holding company, investments are almost entirely financed from government contributions.

6 Conclusion

The aim of this paper was to assess the funding structure of European countries by collecting data that allow to assess differences in the way government support is granted to the railway sector. We used a database of NERA (2004) as starting point of our analysis and developed a taxonomy for the classification of government support to the railway sector. In a second step, we took a deeper look on the funding structure of eight European countries, collected relevant payments and evaluated the consistency of already existing studies. To compare government support between countries we classified payments according to seven categories that reflect the main areas of support as well as the organizational structure of the European railway sector. This involves payments that are reflected as income in the profit and loss statement of infrastructure managers (category 1), payments to support infrastructure investment (2), Public Service Compensations (3), payments for investments in rolling stock (4), pension and staff obligations (5), payments to reduce the indebtedness of undertakings, debt service payments and the takeover of historical debt (6) as well as obligations related to the restructuring of the railway sector (7).

The quality and the coverage of available data on government contributions has increased in recent years. Nevertheless, there are still large gaps in data for some countries. It is clear that due to the complexity of the sector a full coverage of data cannot be achieved. However, no major political attempts have been made to standardize the reporting of government funds being spend on the sector, even though they account for a significant portion of Government budgets. All of the investigated countries support their national railway sector by means of direct contributions that are either granted for the provision of transport services, the operation of the infrastructure as well as to support investments in the infrastructure or rolling stock. The total height of government contributions as well as the breakdown of payments differ to a large extent between countries. Furthermore there are large differences between the degrees to which government contributions cover operating and capital expenditure of infrastructure and transport provision. We compared the development of government support using two main indicators: passenger-ton-kilometer (ptkm) and the number of inhabitants. The average contribution per ptkm amounted to 0.08 PPP adjusted Euro over the observation period (other support than for infrastructure and transport provision excluded). Norway receives the highest contribution per ptkm (0.14 Euro) while Sweden funds every demanded ptkm with an average of only 0.03 Euro. Taxpayer cost (support per inhabitant) amount on average to 140 Euro per inhabitant, with Switzerland receiving the highest average contribution per inhabitant (318 Euro) and Great-Britain the lowest (73 Euro).

Concerning the focus of government support there are also large differences between countries. We identified a tendency to mainly two financing models. In our data-set there are countries that either focus their support payments on the operation of the infrastructure, which implies lower network charges and thus a lower amount of necessary PSC, or they focus on the support of transport services with a higher degree of cost coverage of network charges and thus a lower amount of operating contributions that need to be paid to the infrastructure manager. However, we also find that in some countries no clear tendency can be observed. These countries apply a "hybrid"-like financing approach where the infrastructure manager as well as transport operators receive an almost equal share of operating contributions. Concerning the financing of infrastructure investments two approaches exist. Either investments are financed from government contributions or from funds of the infrastructure manager, which usually involves issuing debt. Some infrastructure managers are compensated for the additional burden of financing debt while other IM need to cover these expenses from network charges and revenue contributions. There is no clear connection between the degree of separation and the choice of a specific financing model observable. However, four of five countries that have completely separated infrastructure and operations are financing infrastructure investments from debt issued by the infrastructure manager. Whereas, in other countries investments are almost entirely financed from government grants.

In 2015 Switzerland has granted the majority of support for infrastructure investment (around 60 percent of total funds) while funds for investments made up only 20 percent in France and 30 percent in Great-Britain. In Germany, Italy, Norway, Spain and Sweden infrastructure investment contributions amount to around 40–50 percent of total support which was granted for infrastructure and transport provision. Germany is the only country in our data set that does not provide revenue contributions to the national infrastructure manager. However, there is an ongoing discussion about the introduction of revenue support related to access charges of freight operating companies. Respective revenue contributions are rather small in France as well as Switzerland, while they account for the majority of total support in Great-Britain. In Norway and Sweden support for infrastructure operation amounted to around 40 percent in 2015. The majority of funds is granted for the provision of transport services in France and Germany, while the shares of Public Service Compensations are comparably small in other countries. There is a negative relationship between funding provided as revenue contributions to the infrastructure and the compensation for the operation of public services, i.e. countries that are supporting public services to a higher degree, grant less funds per ptkm to the infrastructure manager. However, some countries provide funding above the average ratio, which could be an indicator for an inefficient funding structure. This finding needs to be further investigated, especially under consideration of revenue from other sources. It seems likely that the focus of government support has an influence on the performance of the sector, since, e.g. higher track access charges do not simply reduce the size of PSO payments but they also influence the scope for commercial services.

Apart from government funding revenue from passenger and freight transport as well as access charges that are paid for the use of the infrastructure are the main sources of funding. The degree of cost coverage of infrastructure access charges varies between 3 percent (Norway) and 95 percent (Germany). The degree of investments financed from government contributions differs also to a great extent between countries. While in some countries infrastructure investments are financed mainly from government funds (Germany, Norway) other countries finance investments through debt issued by the national infrastructure manager (France, Spain, Great-Britain). This results in increasing debt and additional financial burden. Thus, a company that is indebted to a high degree could be limited in its capacity to act and thus suffer under inefficiencies. This is why researchers argue that a sustainable financing structure can have a large influence on the performance of the sector (see Beck et al. 2013). The database presented here can serve as a starting point for research on the relationship between the volume and structure of public financing on the performance of the railway sector.

Funding: Jan Schäfer thanks Deutsche Bahn AG for funding a PhD fellowship.

References

Arrigo, U. and G. Di Foggia (2013) "Schemes and Levels of State Aid to Rail Industry in Europe: Evidences from a Cross-Country Comparison," European Journal of Business and Economics, 8(3):14–20.

Arrigo, U. and G. Di Foggia (2014) "Policy-Oriented Insights to Enhance Effectiveness and Efficiency of Public Spending: Comparing the Determinants of Subsidies to Railways in Key European Countries," International Journal of Managerial Studies and Research (IJMSR), 2(11):23–35.

Banverket (2008) Swedish Rail Sector Development: Banverket's sector report.

Beck, A., H. Bente and M. Schilling (2013) "Railway Efficiency – An Overview and a Look at Opportunities for Improvement," International Transport Forum, Discussion Paper (2013–2012).

Best, A., B. Görlach, H. van Essen, A. Schroten, U. Becker and R. Gerike (2006) "Size, Structure and Distribution of Transport Subsidies in Europe: A Report to the European Environment Agency," Final Report, 11 December 2006.

Bundesamt für Statistik BFS (2013) Schweizerische Eisenbahnrechnung 2011.

Dehornoy, J. (2011) The Evolution of Public Funding to the Rail Sector in 5 European Countries – A Comparison. In: *Fourth Annual Competition and Regulation in Network Industries (CRNI) Conference*. Brussels.

Deville, X. and F. Verduyn (2012) *Implementation of EU Legislation on Rail Liberalisation in Belgium, France, Germany and The Netherlands: NBB Working Paper No 221*. Brussels: National Bank of Belgium.

ECORYS Nederland BV (2006) Analysis of the financial situation of railway undertakings in the European Union: Report to the European Commission DG-TREN. Rotterdam: ECORYS Nederland BV.

Engel, E., R. Fischer and A. Galetovic (2010) "The Economics of Infrastructure Finance: Public-Private Partnerships Versus Public Provision," EIB Papers, 15(1):40–69.

European Commission (2007) Annexes to the Communication on Monitoring Development of the Rail Market: Commission Staff Working Document, COM (2007) 609. Brussels.

European Commission (2008a) Community Guidelines on State Aid for Railway Undertakings, Communication from the Commission 2008/C 184/07. Official Journal of the European Union, C 184/13. Brussels, 22.07.2008.

European Commission (2008b) Multi-annual Contracts for Rail Infrastructure Quality: Communication from the Commission to the Council and the European Parliament, COM(2008) 54 final. Brussels.

European Commission (2014) Fourth report on monitoring development of the rail market: Report from the Commission to the Council and the European Parliament, COM(2014) 353 final, Brussels, 13.06.2014.

Eurostat (2005) Three decisions on Italy: SCIP, ISPA and Concessionari d'imposta: Euro-Indicators News Release 65/2005. Luxembourg: Eurostat Press Office.

Eurostat (2008) Methodological Treatment of the Capital Injections into RENFE-Operadora in 2006: Letter to Instituto Nacional de Estadistica Madrid. Luxembourg: European Commission, Eurostat Directorate C: National and European Accounts.

Finger, M., Y. Putallaz and D. van de Velde (2015) Solid Infrastructure Financing for Efficient Rail Systems: Results of an Workshop Organized by SNCF Group's Strategy Team.

Instituto Nacional de Estadistica (2007) Methodological Treatment of the Capital Injections into ADIF: Letter from the National Statistics Office to Eurostat. Madrid, Spain.

Laurino, A., F. Ramella and P. Beria (2015) "The Economic Regulation of Railway Networks: A Worldwide Survey," Transportation Research Part A: Policy and Practice, 77:202–212.

Nash, C. (2008) "Passenger Railway Reform in the Last 20 years–European Experience Reconsidered," Research in Transportation Economics, 22(1):61–70.

Nash, C. (2010) "European Rail Reform and Passenger Services – The Next Steps," Research in Transportation Economics, 29(1):204–211.

Nash, C., J.-E. Nilsson and H. Link (2011) Comparing Three Models for Introduction of Competition into Railways – Is a Big Wolf so Bad After All? CTS Working Paper 2011:19, Centre for Transport Studies Stockholm.

NERA (2004) Study of the Financing of and Public Budget Contributions to Railways: A Final Report for European Commission, DG TREN. London: NERA National Economic Research Associates.

NERA (2005) Public Budget Contributions to the Railways: A Report to Deutsche Bahn AG. London: NERA National Economic Research Associates.

OECD (2010) The PSE Manual – OECD's Producer Support Estimate and Related Indicators of Agricultural Support. Paris: OECD Trade and Agriculture Directorat.

Perkins, S. (2005) The Role of Government in European Railway Investment and Funding. China Railway Investment and Finance Reform Forum, Beijing, China.

RGL Forensics, Frontier Economics and Aecom (2009) Study on Separation of Accounts of Railway Undertakings and Rail Infrastructure Managers: Final Report for the European Commission.

Steer Davis Gleave (2005) RAILIMPLEMENT – Implementation of EU Directives 2001/12/EC, 2001/13/EC and 2001/14/EC: Final Report for the European Commission, Directorate General for Energy and Transport. Steer Davies Gleave, Bologna, Italy.

Valsecchi, C., P. ten Brink, S. Bassi, S. Withana, M. Lewis, A. Best, F. Oosterhuis, C. Dias Soares, H. Rogers-Ganter and T. Kapenhengst (2009) Environmentally Harmful Subsidies: Identification and Assesment: Final report for the European Commission's DG Environment. Institute for European Environmental Policy (IEEP), London and Brussels.

van de Velde, D., C. Nash, A. Smith, F. Mizutani, S. Uranishi, M. Lijesen and F. Zschoche (2012) EVES-Rail, Economic Effects of Vertical Separation in the Railway Sector: Full Technical Report for CER, Community of European Railways and Infrastructure Companies, by inno-V (Amsterdam) in Cooperation with University of Leeds – ITS, Kobe University, VU Amsterdam University and Civity Management Consultants.

A.2. Financing railways

Authors: Georg Götz and Jan Thomas Schäfer

Published in: *Handbook on Railway Regulation*, Edward Elgar Publishing, 2020, pp. 304-321.

20. Financing railways

Georg Götz and Jan Thomas Schäfer

20.1 INTRODUCTION

Financing of railways differs greatly among countries of the European Union. This complicates the comparability of network charges, cost recovery rates, and ultimately the comparison of the financial performance across countries. In this chapter, we will outline differences in the financing structure of railways and its impact on current regulatory matters.

Prior to the liberalization of European Railways in the 1990s, most railways were organized as government institutions, with full integration of infrastructure and operations. Accordingly, services were dependent on state budgets. Governments have required tariffs for public services being determined by political considerations rather than from a business perspective. Due to low profitability, high investments, and tight budgets, many state railways accumulated huge debts (see ECMT 1998, pp. 17–18 and ECMT 2001, p. 45) and government grants were often used to pick up losses rather than being targeted to specific activities or projects. Therefore, one of the goals of Directive 91/440/EEC (European Commission 1991) was to improve the financial conditions of Railways in Europe, while fostering the efficient use of funds. The (legal) separation of infrastructure and operations and the introduction of access charges have had a significant impact on the financing framework. Even though infrastructure and operations are financed independently, there remains a strong financial dependence, with access charges determining income and expenses of infrastructure managers and train operating companies, respectively.

More than 20 years after the reform of the sector, the (financial) involvement of governments remains high and some infrastructure managers have accumulated large debts. The structure of public financing to European infrastructure managers and railway undertakings has been studied in several publications. For a recent contribution, see Schäfer and Götz (2017). In the wake of initiatives to increase the modal shares of eco-friendly transport, it is becoming clearer that governmental influence will remain at this high level or even continue to increase in future. From a regulatory point of view, this means that control of the efficient use of public funds is still an important topic, particularly regarding payments to the monopolistic infrastructure. In terms of public financing of transport operations, efficiency might benefit from the application of tendering (European Conference of Ministers of Transport 2007; Nash & Wolański 2010). However, as Bajari et al. (2009) showed, competitive tendering might also suffer from limitations and it is unsurprising that some share of services is still being awarded directly.

Standard literature on the regulation of natural monopolies and the determination of access charges has identified marginal or incremental-cost charging as being optimal. However, to cover the total costs of the infrastructure, transfers are required. If transfers are costly or budget restrictions apply, different second-best solutions are feasible. The wording in Directive 2012/34/EU (European Commission 2012) points towards the application of Ramsey–Boiteux pricing with access charges being determined on the level of short run marginal-cost plus "what the market can bear" (Nash 2017). However, the market is in large part organized and financed by the government itself; that is, a fraction of the expenses of train operating companies is covered by public service contributions rather than by market revenue. In practice, an increase in the cost recovery of the infrastructure only leads to a shift in funding. This raises the question of whether pooling of public funding at the infrastructure level to keep network charges close to marginal costs could be beneficial. Sánchez-Borràs et al. (2010) pointed out that commercial transport in particular could benefit from lower mark-ups.

Public funding regularly causes incentive and efficiency problems. The coexistence of contractual agreements between the government and the infrastructure manager, the simultaneous regulation of access charges, and the occurrence of the government as a customer of transport services make it difficult to optimize the efficiency of the railway sector as a whole. This is particularly the case when parts of the costs covered by public financing are exempted from regulation. On the other hand, contractual public funding has a positive impact on financial stability of infrastructure managers. It seems important that the financing of investments, renewals, and maintenance are secured in the long term. While multi-annual contracts seem to improve financial sustainability, it is important that contracts also contain targets whose achievement is monitored and sanctioned (Cowie 2009). Recent developments show that regulators are also becoming increasingly involved in all stages of the investment lifecycle (IRG-rail 2018).

The remainder of this chapter is structured as follows. Section 20.2 summarizes the EU legislation on railway financing. The financing practice of infrastructure development and management, as well as the financing of the operation of transport services, will be discussed in sections 20.3 and 20.4, respectively. Section 20.5 concludes by summarizing the main challenges when it comes to regulation and regulatory policies under strong involvement of government funds.

20.2 THE EU FINANCING FRAMEWORK

Generally, a distinction can be made between two types of financing: market-based financing and public financing. The latter is an important pillar of railway financing in Europe. Unlike other sectors, public financing to infrastructure managers and railway undertakings can be easily declared compatible with the European Treaties if they meet the needs of coordination of transport or if they represent reimbursement for the discharge of certain obligations inherent in the concept of a public service (see Article 93 of the Treaty on the Functioning of the European Union). This leaves room for different types of public financing, with the broad framework specified by further regulation. The financing framework is different for the infrastructure manager and transport undertakings, yet both financially depend on each other.

20.2.1 Infrastructure

Regarding financing of the infrastructure, it is important to distinguish between two activities: the management and operation of the railway infrastructure, including maintenance, and the renewal and development of the infrastructure, which is investments to upgrade and expand the existing network. According to Directive 2012/34/EU, access charges (that is, market-based revenue), must at least cover the cost that is directly incurred as a result of operating train services on the network. Direct costs may include maintenance, renewal, and operating costs (Nash et al. 2018). In practice, the exact allocation and determination of direct cost is difficult. Revenue from access charges is hardly ever enough to cover fixed costs, including investment expenditure. This is why the aforementioned Directive permits mark-ups on the cost directly incurred, in order to obtain full cost recovery. However, no major infrastructure manager in the European Union has achieved full cost recovery (see section 20.3). The level of cost recovery through infrastructure charges affects the necessary level of government contributions. Also, Member States may require different levels of overall cost recovery; that is, the level of charges can differ greatly between countries.

According to Directive 2012/34/EU, governments have to ensure that the profit and loss accounts of the infrastructure managers are balanced. Specifically, revenue from access charges, surpluses from other commercial activities, and non-refundable incomes from private sources and government contributions must cover all infrastructure expenses that, according to international accounting rules, appear in the income statement of the infrastructure manager. This typically involves operating expenses such as maintenance and repairs, materials, salary, and wages, as well as depreciation for non-grant-funded investment. In contrast, capital expenditure for renewals and the development of the network do not directly show up in the income statement. Directive 2012/34/EU points out that Member States may decide to finance these investments through means other than direct government contributions. Therefore, different approaches may be considered for the financing of investments. In some countries, investments are financed almost exclusively from government funds, while other countries finance infrastructure projects from funds that infrastructure managers or independent financing entities raise on the financial markets.

The problem of cross-subsidization is one of the major concerns regarding public financing of railways. This is why Directive 2016/2370/EU (European Commission 2016a), amending Directive 2012/34/EU, states that financial transfers between the infrastructure manager and railway undertakings, and in vertically integrated undertakings between the infrastructure manager and any other legal entity of an integrated undertaking, should be prevented if they could lead to a distortion of competition on the market, particularly as a result of cross-subsidization. Directive 2012/34/EU also prohibits cross-subsidization between the accounts of passenger and freight transport services, as well as between services that are and are not subject to a public-service remit.

20.2.2 Transport

Regarding the financing of railway undertakings, one must distinguish between services that are subject to a Public Service Obligation and other services. Whereas public services usually involve public financing, other services typically run without government support and finance operation entirely from market revenue. However, under certain circumstances, European legislation also permits public contributions to such services that are operated commercially. This refers to investment contributions, such as rolling stock, rather than to operating contributions.

The scheme for the compensation of public passenger transport services by rail was laid down in Regulation 1370/2007 (European Commission 2007), which was recently amended by Regulation 2016/2338 (European Commission 2016b). Payments that are granted to ensure the provision of services which are of general interest may not exceed what is necessary to cover the net costs incurred through discharging the public service obligations, taking account of the revenue generated thereby and a reasonable profit. The regulation does not specify the form in which public service compensations must be granted. It refers to any benefit, especially financial benefit, that might be granted, either directly or indirectly. However, the method of compensation must promote the maintenance and development of effective management and the provision of passenger transport services of a sufficiently high standard. In practice, government contributions to railway undertakings, which are operating public services, are granted as a substitute for market revenue and are therefore recognized as income in the profit-and-loss statements of train operating companies. In some cases, investments in rolling stock are compensated separately or organizing authorities purchase rolling stock themselves and lease them to railway undertakings.

Government payments to railway undertakings that do not provide public services can be declared compatible with the Treaties under certain circumstances, as described in the Commission Guidelines on State aid for railway undertakings (European Commission 2008). The guidelines cover several fields: infrastructure investments, renewal of rolling stock, debt cancellation, restructuring, coordination of transport, and guarantees for railway undertakings. Aid that falls under the field of coordination of transport has recently been applied in several Member States, such as Austria and Germany; that is, aid that aims to lower the cost for rail infrastructure use and aid to reduce external cost of freight services (see section 20.4). In contrast to most financing instruments mentioned above, aid to railway undertakings that are not subject to a public service obligation might require legal approval.

20.3 INFRASTRUCTURE FINANCING IN PRACTICE

In this section, we will analyze the financing structure of 12 European infrastructure managers. Investment spending and financing of capital expenditure will be analyzed in section 20.3.1, while the financing of the management and operation of the infrastructure will be studied in section 20.3.2. The data set includes infrastructure managers with different network sizes and organizational structures. All studied infrastructure managers are 100 percent state-owned and operate as independent entities or as part of a holding structure. In some countries, infrastructure managers are organized as government agencies rather then commercial companies. The Spanish infrastructure is managed by two companies: ADIF manages the conventional network, while ADIF AV (Alta velocidad) is responsible for the Spanish high-speed network. The state-owned enterprise Bane NOR was founded in 2017 and is the successor of Jernbaneverket, the former Norwegian railway agency. DB Netz is the German state-owned infrastructure manager that operates the largest network in Europe. German Railways (DB) is organized in a holding structure. Infrabel is responsible for the Belgian railway network. Since 2014, Infrabel has been an autonomous public company (no longer owned by the SNCB-Holding). The British rail infrastructure is owned and managed by Network Rail, a government-owned company. ÖBB Infra is responsible for the management and construction of the Austrian rail infrastructure. The Dutch railway infrastructure is managed by a government task organization called ProRail. RFI is a subsidiary of the Italian Ferrovie dello Stato (FS) Holding and owns and maintains the Italian Network. SBB Infrastructure is responsible for the management of the Swiss railway infrastructure. SNCF Réseau is the successor of the former independent RFF. SNCF Réseau is part of the SNCF Holding structure. The Swedish

Transport Administration, Trafikverket, is a government agency that is responsible for road, rail, shipping, and aviation.

20.3.1 Renewal and Development of Infrastructure

Infrastructure investments constitute a major part of the expenses of the whole railway sector (Casullo 2017). There are basically three different sources of funds to finance infrastructure investments: (1) government grants, (2) funds from the financial market (loans or bonds), and (3) own funds (retained earnings/cash or equity). Loans can refer to funds borrowed from governments or from the financial markets. In some countries, special purpose entities are responsible for financing infrastructure investment. For example, the Swiss Rail Infrastructure Fund (BIF) finances the operation and development of Swiss railways. Some special purpose entities manage tax revenues, as in Switzerland, and/or raise loans or issue bonds. Financing from the financial markets can result in further grants; that is, support to repay debt. This is why the debt of state-owned railway managers is usually included in the public debt account (Maastricht criteria). Governments sometimes transfer funds for investments by the means of equity contributions.

To be able to compare the financing of infrastructure investment, it is crucial to understand the differences between the three above-mentioned sources of financing. The way in which infrastructure investments are financed has a direct impact on the level of operating expenditures − that is, depreciation and interest expenses − so it also influences the amount of funds that need to be generated in order to repay loans. In Great Britain and France, for example, investments are financed mainly from (government) loans. Therefore, Network Rail and SNCF Réseau need to cover 2 billion euros worth of depreciation as well as 2.5 billion euros of interest expenses from access charges and other operating income, including operating grants (see section 20.3.2). Although the German infrastructure manager DB Netz invested a similar amount, depreciation is just 0.8 billion euros and interest expenses are as low as 0.2 billion euros. However, almost all investments in Germany are financed from government grants.

Our analysis essentially builds on the accounting treatment of grants, income, and expenditure. In this chapter we focus on payments that are directly related to the financing of infrastructure investment; that is, renewals and the development. Expenses that are indirectly related to infrastructure investment (such as depreciation and interests) are accounted for in the next section. According to international accounting standards (such as IAS 20, Accounting for Government Grants, and Disclosure of Government Assistance), government grants should be recognized in profit-or-loss statements on a systematic basis over the periods in which the entity recognizes expenses for the related costs for which the grants are intended to compensate. This can be done either by setting up the grant as deferred income or by deducting it from the carrying amount of the asset. Both means lead to lower periodic expenses that the IM needs to cover from own resources; that is, infrastructure managers do not have to cover for depreciation of grant-funded investments (see Germany above).

In contrast, if the government decides to let the infrastructure manager raise funds on the financial markets rather than to finance investments from direct investment contributions, the IM needs to cover depreciation and interest expenses from income. The same holds for investments that are financed from the own funds of the infrastructure manager (see Great Britain and France above). In most cases, market revenue is not enough to cover these expenses, which is why additional grants are often transferred to the IM. If infrastructure managers receive compensation for investment-related expenditures (depreciation, interest expenses) they are typically accounted for in the income statement and will be studied in section 20.3.2. For example, the Austrian infrastructure manager ÖBB Infra finances a major part of investments from funds borrowed from the government (Österreichische Bundesfinanzierungsagentur), but in return receives an annual contribution to cover interest expenses and to repay a part of its debt.

Table 20.1 contains data on investment spending, investment grants and the level of liabilities of European infrastructure managers in 2018. It also indicates whether or not an infrastructure manager borrows funds from the financial markets or the government. Investments comprise renewals and expansion of the network. Figures in parentheses indicate change from the previous year. For an evolution of grants over time, see Schäfer and Götz (2017). All infrastructure managers receive government contributions (investment grants) to finance investments. For better comparability, Table 20.2 shows investments and grants per inhabitant. Table 20.3 shows the percentage of infrastructure investments that are financed from direct government contributions. Other funding sources comprise borrowings from the financial markets and own funds of infrastructure managers. If companies rely on loans to finance investment, liabilities react accordingly. Some infrastructure managers have accumulated large debt accounts amounting up to 70 billion euros (Network Rail and SNCF Réseau).

Main sources of funding for ADIF-AV include investment grants from national, regional, and EU funds. Furthermore, ADIF-AV borrows funds from the financial markets, up to the annual limit set by the General State Budget Laws for each year. ADIF applies a similar approach of financing. However, investments of ADIF are almost 90 percent covered by government grants, and ADIF AV finances 70 percent of investments from funds other than grants. In 2018, ADIF and ADIF-AV invested 36.59 euros per inhabitant, of which 16.59 euros per inhabitant was financed from government grants. At the same time, liabilities of ADIF-AV increased by 0.63 billion euros.

Nearly all investments of Bane NOR are financed from government contributions. In 2018, Bane NOR invested 1.4 billion euros and received the same amount in investment grants. This corresponds to an annual investment of 264.18 euros per inhabitant, which is the largest value among the countries studied (note that figures have not been converted using purchasing power parity).

In Germany, investments of DB Netz (renewals and development) are largely financed by government grants. About 10 percent of investments are financed from the infrastructure manager's own funds. Investments and investment grants have been increasing from previous years. In 2018, DB Netz invested nearly 6.9 billion euros and received 6.3 billion euros in investment grants. Grants for renewals are subject to a multi-annual contract. Spending per inhabitant reached 83.24 euros in 2018.

The Belgian infrastructure manager Infrabel finances all investments from government contributions. In 2018, 77.64 euros was spent per inhabitant. Investments are subject to a multi-annual investment plan with the government. The regions put in co-financing for carrying out certain priority projects.

In Great Britain investments are mainly financed from government loans and a minor part from investment grants. A loan facility agreement between the Secretary of State for Transport and Network Rail covers Network Rail's financing requirements for the current control period. In 2018, Network Rail invested 7.5 billion euros (113.12 euros per inhabitant). Liabilities rose by 4 billion euros. Network Rail also received investment grants of 0.9 billion euros (13.49

Table 20.1 Financing of infrastructure investment

	ADIF	**ADIF Alta velocidad**	**Bane NOR**	**DB Netz**	**Infrabel**	**Network Rail**	**ÖBB Infra**	**ProRail**	**RFI**	**SBB Infrastructure**	**SNCF Réseau**	**Trafikverket**
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Network length in km	11,992	3,153	3,848	33,299	3,602	15,878	4,864	3,075	16,781	3,228	30,013	10,906
Investment spending	397 (+1)	1,310 (-36)	1,399(+93)	6,892 (+302)	885 (+57)	7,497 (-174)	1,985 (+177)	957 (+28)	4,549 (+340)	2,051 (+54)	5,057 (-321)	979 (-44)
Investment grants	346 (−158)	428 (+185)	1,399 (+143)	6,338 (+397)	885 (+57)	894 (+401)	120 (−8)	902 (+41)	4,681 (+548)	50 (−14)	1,546 (−216)	847 (−64)
Loans	Yes	Yes				Yes	Yes			Yes	Yes	Yes
Liabilities	4,770 (+52)	21,472 (+627)	16,792 (+919)	10,710 (+150)	4,998 (-96)	70,162 (+4,038)	22,739 (+922)	17,461 (+355)	12,178 (+91)	18,970 (+707)	66,665 (+3,108)	n/a

Note: Figures reported in millions of euro. NOK, GBP, CHF, and SEK converted using annual average exchange rates. Figures in parentheses stand for change from previous year.
Source: Annual reports of infrastructure managers.

Table 20.2 *Investments and grants per inhabitant in euros*

	ADIF	ADIF Alta velocidad	Bane NOR	DB Netz	Infrabel	Network Rail	ÖBB Infra	ProRail	RFI	SBB Infrastructure	SNCF Réseau	Trafikverket
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Investments	8.51	28.08	264.18	83.24	77.64	113.12	225.00	55.70	75.21	241.75	75.56	96.74
Investment grants	7.42	9.17	264.18	76.55	77.64	13.49	13.60	52.50	77.39	5.89	23.10	83.69

Note: NOK, GBP, CHF, and SEK converted using annual average exchange rates.

Table 20.3 *Percentage of investment financed from government grants*

	ADIF	ADIF Alta velocidad	Bane NOR	DB Netz	Infrabel	Network Rail	ÖBB Infra	ProRail	RFI	SBB Infrastructure	SNCF Réseau	Trafikverket
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Grant ratio	87.15	32.67	100.00	91.96	100.00	11.92	6.05	94.25	102.90	2.44	30.57	86.52

euros per inhabitant). The remainder of investments are financed from retained earnings, although Network Rail receives a operating grant that accounts for a major part of operating income (see next section).

As noted earlier, ÖBB Infra's investments are financed with funds borrowed from the government. Accordingly, ÖBB Infra's liabilities increased by 1 billion euros. Additionally, ÖBB Infra received capital grants and finances a part of investments from the cash flow from operating activities. Investment per inhabitant equaled 225 euros, 6 percent of which was financed from investment grants.

In the Netherlands investments are financed by the Ministry of Infrastructure and Water Management and other authorities. In 2018, ProRail invested 0.95 billion euros, all of which was financed from government grants. ProRail's investments amounted to 52.50 euros per inhabitant.

Italy uses an approach whereby investments are nearly entirely financed from investment grants. In 2018, RFI invested 4.5 billion euros and received 4.7 billion of investment grants from the central government, the EU, and local bodies. The surplus of grants is not commented on in the annual report of RFI, but could be due to different demarcation of investments and grants. Investments amounted to 75.21 euros per inhabitant.

Investments for renewals and expansion of the Swiss SBB network are financed from government loans and non-repayable contributions (BIF fund). SBB also uses own funds from the cash flow to invest. A small fraction of these funds come from market revenue and a large part come from a operating compensation that SBB receives to cover for depreciation. The compensation amounted to 1.4 billion euros in 2018 (covering more than 60 percent of investment). Since depreciation does not require a current outlay of cash, SBB can use these funds to invest. The depreciation compensation is treated as income in SBB's profit-and-loss statement, even if it is primarily used to finance investment. Thus, payments are treated as operating grants, which will be studied in section 20.3.2. Approximately 40 percent of investments are financed from funds borrowed from the government. Non-repayable grants and own funds account for around 50 million Swiss francs each. In 2018, SBB Infrastructure spend was 241.75 euros per inhabitant for renewals and development.

The French SNCF Réseau raises funds on international capital markets in order to cover its long-term financing needs. This is mainly done through bond issues and private placements and, to a small extent, by borrowing from banks or local authorities. The French infrastructure manager also receives funds from the French state, the Agence de financement des infrastructures de transport de France (AFITF) and other regional authorities. In 2018, SNCF Réseau invested 75.56 euros per inhabitant, 30 percent of which were received as investment grants.

The Swedish infrastructure manager Trafikverket performs tasks for different modes. Some accounts are shared between modes, which makes it difficult to determine the cost and funding structure of rail infrastructure. The investments to be undertaken by Trafikverket are determined in the National Transport System Plan and the county plans for regional infrastructure. Funding is provided accordingly. Furthermore, Trafikverket receives TEN-T funding from the European Union and covers a small fraction of investments from railway charges. In the past, Trafikverket has borrowed funds from the Swedish National Debt Office and continues to receive compensation to repay debt and to cover interest expenses. In 2018, Trafikverket invested the equivalent of 96.74 euros per inhabitant, 83.69 euros of which was financed from investment grants related to railways.

The comparison of different approaches of investment financing shows that public funds play a major role, especially for enhancements and partially for renewals. The public share is also high in countries where the government allows the infrastructure manager to raise funds on the financial markets rather than to finance investments from direct investment contributions, because the deficit can or is often treated as state deficit that the IM will likely never be able to manage without additional public funds that are transferred as operating contributions. The Maastricht constraint will most likely lead to changes in investment financing, since railway debt is now treated as public debt. In Great Britain, Network Rail has stopped borrowing funds from the financial markets and investments are now funded from government loans. France will restructure the financing of investments of SNCF Réseau in 2020. The way investments are financed and accounted for in the accounts of the infrastructure managers has an direct impact on operating expenses that have to be covered from access charges and other revenue.

20.3.2 Infrastructure Management

According to Directive EU 2012/34, revenue from access charges, surpluses from other commercial activities, non-refundable incomes from private sources, and government contributions must balance the profit and loss accounts of infrastructure managers over a reasonable time period. Table 20.4 shows key statistics from the income statements of infrastructure managers as well as the number of train kilometers run on the network in 2018. Operating income typically includes revenue from access charges and operating grants. In some countries, investments grants are deducted from the carrying amount of the asset, while other infrastructure managers release investment grants to income. The latter option has a positive impact on the reported operating income, with higher depreciation as a counter position. Some IMs, such as the French SNCF Réseau, have large positions of own work capitalized in their operating income. This is to balance production cost of additions to fixed assets that have been produced by the company itself and have led to material, labor, and other production costs. Operating expenses typically include staff and depreciation, as well as spending for material and services. The amount of loans to finance infrastructure projects does not appear in profit-and-loss accounts, but depreciation and interest expenses do.

Except for the two Spanish infrastructure managers, all IMs receive operating grants. Table 20.5 shows the cost coverage ratios of total non-grant income (that is, operating income less government grants). Only infrastructure managers in Spain, Germany, and France reach full operating cost recovery from non-grant income. However, after accounting for interest payments (debt service), the Spanish and French infrastructure managers incur losses due to the large expenses that companies need to bear to cover cost of infrastructure financing. The operating grants paid to Network Rail, ÖBB Infra, and SBB Infrastructure include fractions that are dedicated to finance parts of investments or related expenses, such as deprecation and interests. The operating grant to Network Rail is not allocated to cover particular expenditure, although the government specifies the level of network enhancements (IRG-rail 2018).

Both Great Britain and Sweden aim to set access charges at a marginal-cost level (Nash et al. 2018). Network Rail charges a mark-up to specific freight transport services and has a lump-sum mark-up for franchise operators in place. To make figures comparable, Table 20.6 shows the average revenue from track access charges per train kilometer as well as the amount of operating contributions per train kilometer. The level of track access charges varies between countries. Typically, countries with low access charges receive higher operating grants.

Table 20.4 Financing of infrastructure operation and maintenance

	ADIF	ADIF Alta velocidad	Bane NOR	DB Netz	Infrabel	Network Rail	ÖBB Infra	ProRail	RFI	SBB Infrastructure	SNCF Réseau	Trafikverket
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Million train kilometers	139	58	47	1,071	110	553	154	164	363	182	441	159
Operating income	1601	1,101	1,326	7,373	1,442	9,472	3,304	1,781	3,844	3,829	12,488	n/a
access charges	617	473	35	5,105	760	1,929	535	340	1,175	1,058	5,275	172
grants	0	0	534	416	108	5,063	1,877	745	1,016	1,742	164	908
Operating expenses	1,673	1,052	1,325	6,667	2,109	7,171	2,701	1,786	3,531	3,886	11,969	n/a
staff	631	14	286	2,750	692	2,734	1,183	284	1,497	1,341	2,347	n/a
depreciation	412	345	614	760	697	1,851	777	590	106	1,181	2,278	n/a
EBIT	-72	49	2	706	-667	2304	603	-5	312	-57	517	n/a
Net financial result	-3	-273	28	-209	704	-2,514	-558	-7	-38	n/a	-1,484	n/a
interest payments	13	359	7	210	61	2,523	579	8	38	n/a	2,363	n/a
EBT	-74	-223	30	511	36	54	45	-12	274	n/a	-967	n/a

Note: Figures reported in millions of euros. NOK, GBP, CHF, and SEK converted using annual average exchange rates.
Source: Annual reports of infrastructure managers.

Table 20.5 *Cost coverage of non-grant income in percent (operating income less operating grants divided by operating expenses, including depreciation)*

	ADIF	**ADIF Alta velocidad**	**Bane NOR**	**DB Netz**	**Infrabel**	**Network Rail**	**ÖBB Infra**	**ProRail**	**RFI**	**SBB Infrastructure**	**SNCF Réseau**	**Trafikverket**
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Cost coverage ratio	95.70	104.66	59.79	104.35	63.25	61.49	52.83	57.91	80.09	53.71	102.97	n/a

Table 20.6 *Track access charges and operating grants in euros per train kilometer*

	ADIF	**ADIF Alta velocidad**	**Bane NOR**	**DB Netz**	**Infrabel**	**Network Rail**	**ÖBB Infra**	**ProRail**	**RFI**	**SBB Infrastructure**	**SNCF Réseau**	**Trafikverket**
Country	ES	ES	NO	DE	BE	GB	AT	NL	IT	CH	FR	SE
Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Track access charges	4.43	8.15	0.74	4.77	6.92	3.49	3.47	2.07	3.24	5.80	11.96	1.08
Operating grants	0.00	0.00	11.30	0.39	0.98	9.15	12.16	4.54	2.80	9.55	0.37	5.72

Note: NOK, GBP, CHF, and SEK converted using annual average exchange rates.

The average revenue from access charges is only 0.74 euros per train kilometer in Norway, while the French SNCF Réseau redeems almost 12 euros per train kilometer. Accordingly, the Norwegian infrastructure manager Bane NOR receives a compensation of 11.30 euros per train kilometer, while the French infrastructure manager receives operating grants of only 0.37 euros per train kilometer.

The large difference between the sum of track access charges and operating grants per train kilometer can be explained by different cost structures. For example, SNCF Réseau and ADIF-AV operate a large high-speed network, while Trafikverket is organized as an government agency that shares competencies with other sectors like road and air. Differences can also be explained by different price levels.

20.4 TRANSPORT FINANCING IN PRACTICE

Track access charges are among the main expenditures of train operating companies. In Germany, they account, on average, for 20−40 percent of cost of railway undertakings (Bundesnetzagentur 2013). While the analysis of the financing structure of infrastructure managers is relatively easy, the analysis of accounts of train operating companies requires a coarser approach. This is for several reasons. In many countries the market has been opened for competition. This is particularly true for freight transport, but also for commercial passenger transport and for public services. Accordingly, annual reports of several hundred companies would have to be investigated. Since the former state-owned enterprises still retain a major part of market shares, we have focused our analysis on these operators. Unfortunately, it is not possible to compare accounts of different divisions directly as not all operators publish accounts that differentiate between commercial services and services that are subject to a public service obligation (PSO). However, these divisions are of particular interest from a regulatory point of view since the application of mark-ups is particularly contentious if costs are somehow borne by the government.

While long-distance services and freight services typically run without government grants (that is, their operation is financed entirely from market revenue), regional services receive compensation that substitutes for market revenue. In some countries some medium and long distances are also run under a PSO (France, Austria, Sweden). PSO services account for the major part of traffic on the network in most countries. This is certainly true in terms of the number of passengers, but also for the number of train kilometers (see, for example, the train kilometers in the PSO and commercial segment, respectively, in Spain and Germany in Table 20.7). Thus, they are the main customers of infrastructure managers. It is likely that some parts of expenses for the use of the infrastructure are covered by public funds that train operating companies receive from the government. In France, for example, the organizing authorities pay the charges for the use of the infrastructure directly to SNCF Réseau rather than transfer these funds to the train operating companies. In countries with high degrees of infrastructure cost coverage, some parts of the cost for the use of the infrastructure are passed on to the organizing authorities. Thus, the determination of mark-ups permitted by Directive 2012/34/EU seems to be critical for PSO services. Estimation of the elasticity of demand of the government in order to apply Ramsey–Boiteux pricing appears a demanding task. Still, most countries also apply mark-ups on the track access charges of PSO services (IRG-rail 2018).

Table 20.7 Market revenue, public service compensation and track access charges in million euros

	Renfe Operadora	Renfe Operadora	NBS AS	DB Regio	DB Fernverkehr	SNCB	SNCB	Franchises	ÖBB	Trenitalia	SBB	SNCF Mobilities
Country	ES	ES	NO	DE	DE	BE	BE	GB	AT	IT	CH	FR
Year	2017	2017	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Segment	PSO	COM	ALL	PSO	COM	PSO	COM	PSO	ALL	ALL	PSO	ALL
Million train kilometers	107	61	39	380	142	n/a	n/a	516	97	247	80	n/a
Market revenue	750	1,419	430	2,865	4,448	1,038	272	10,885	977	2,962	838	8,405
PSO	864	0	366	3,419	0	1,216	0	2,536	1,110	2,040	566	5,251
net								−450				
Access charges	413	506	n/a	1944	987	n/a	n/a	n/a	348	846	n/a	3828

Note: COM refers to commercial service segments, ALL includes public and commercial passenger transport activities. NOK, GBP, and CHF converted using annual average exchange rates.

Table 20.8 Market revenue, public service compensation and track access charges in euros per train kilometer

	Renfe Operadora	Renfe Operadora	NBS AS	DB Regio	DB Fernverkehr	SNCB	SNCB	Franchises	ÖBB	Trenitalia	SBB	SNCF Mobilities
Country	ES	ES	NO	DE	DE	BE	BE	GB	AT	IT	CH	FR
Year	2017	2017	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Segment	PSO	COM	ALL	PSO	COM	PSO	COM	PSO	ALL	ALL	PSO	ALL
Market revenue	6.99	23.29	10.99	7.54	31.28	n/a	n/a	21.11	10.03	11.98	10.42	n/a
PSO	8.05	0	9.35	9.00	0	n/a	n/a	5.56	11.40	8.25	7.04	n/a
net								−0.77				
TAC	3.85	8.31	n/a	5.12	6.94	n/a	n/a	n/a	3.57	3.42	n/a	n/a

Note: COM refers to commercial service segments, ALL includes public and commercial passenger transport activities. NOK, GBP, and CHF converted using annual average exchange rates.

Table 20.7 shows the public service compensation that is paid to the main train operating companies. Detailed information on the cost and expenses of PSO services is only available for Spain and Germany; no other companies publish more detailed accounts of PSO and Non-PSO services separately. In Spain and Germany, about 30 percent of operating income is used to cover expenses related to the use of the network (track access charges). Train operating companies in Great Britain have been privatized. The government tenders franchises for profitable and non-profitable services. In sum, a positive revenue stream for the British government is generated. Table 20.8 compares the public service compensation per train kilometer. Compensation ranges from -0.77 euros in Great Britain to 11.40 euros per inhabitant in Austria. However, to complete the analysis one would also have to compare ticket prices of public services.

The ability of train operating companies to bear mark-ups is influenced not only by the amount of government support that is granted, but mostly by the competitive environment. Recent cases have shown that freight rail services are not always able to bear high mark-ups. Germany and Austria have applied for approval of state aid to freight rail operators. In Germany these payments are intended to lower the cost of infrastructure access of operators and amount to 350 million euros per year (European Commission 2019). This equates to a reduction of access charges of 1.30 euros per train kilometer. Austria set up an annual budget of 120 million euros to support rail freight undertakings (European Commission 2018). In 2018, 41 million freight train km have been run on the ÖBB Infra network; state aid amounted to around 2.90 euros per train kilometer. In Great Britain, Switzerland, France and other countries, infrastructure managers or freight rail operators receive compensation in order to lower cost and shift transport to rail. The Europen Commission's State Aid Database lists cases for many more countries that have applied for approval of grants related to the coordination of freight transport.

20.5 CONCLUSION

Infrastructure regulation is moving towards the implementation of Ramsey–Boiteux-like access charges. Even though user financing is an important tool with which to minimize the social cost of public funds, experience shows that the level of government funding remains high. It is not clear whether an increase of access charges in order to achieve higher cost recovery ratios will lead to an increase in efficiency, or even a reduction of total public support to the sector. A deviation from marginal-cost pricing of infrastructure access is likely to lead to a reduction in rail transportation. Today, more than 80 percent of train kilometers (65 percent of passenger kilometers) in the European Union are organized under PSOs (IRG-rail 2017, 2019). These services are typically not profitable and receive government contributions; an increase of access charges can only partly be passed on to consumers while a majority remains with the government. At the same time, commercial services that face strong intermodal competition will become less profitable.

The case of rail freight transport has shown that some commercial services can only compete with other modes if they receive additional public financing. This might contradict the goals formulated in the EU White Paper on the Roadmap to a Single European Transport Area (European Commission 2011). Catalano et al. (2019) argued that it is necessary to move away from traditional technical or allocative efficiency analyses and use a broader approach that considers transport policy goals in order to optimize the sector as a whole. When regulating infrastructure charges, the level of government financing on all stages should be taken into account as well as the degree to which all modes of transport are priced according to their respective environmental cost. Accordingly, and even though the first railway liberalization packages aimed to move away from or reduce public financing by achieving full cost recovery, one should expect that this will not be the case for the next few decades.

Given that policy involvement will remain high and that significant public funds are likely to finance investments and operation of network infrastructure, it is important to use funds efficiently. In most countries, access charge regulation and multi-annual contracts that determine the financing framework of the infrastructure manager are independent, at least to some extent. This can cause inefficiencies if government-funded expenditures are exempted from access-charge regulation, if contractual agreements between the government and the infrastructure manager do not offer sufficient incentives to use funds efficient, or if access regulation and multi-annual contracts provide contradictory incentives. It seems preferable to use more integrated approaches that combine the regulation of charges, as well as schemes to promote efficient use of public funds.

A more integrated approach would also mean that regulators should consider government financing on all stages of the railway sector, not only funds that are granted to the infrastructure manager. For example, it is difficult to determine the elasticity of demand of railway services that receive government funding since mark-ups are partly borne by the governments rather than by consumers. Therefore, one might exempt these segments from bearing mark-ups, at least as long as network expansion is not at issue. For other segments, mark-ups might be more feasible. However, when determining access charges and mark-ups, one must take account of environmental targets and differences between the cost structures of modes. It seems appropriate to reallocate a fraction of public service contributions to the infrastructure manager in order to lower the overall level of access charges.

With regard to the plans to establish a Single European Railway Area, different approaches to infrastructure financing and large differences in the level of access charges seem to have harmed the development of the sector. This is particularly true for cross-border freight in the EU, which also suffers from heterogeneous technical requirements (Zunder et al. 2013). Although the latest EU directives aim to harmonize access charges throughout Europe, it turns out that the level of infrastructure charges is still far from being uniform. This is mainly due to the different financing approaches (different requirements of cost coverage of charges and different levels of public funding) rather than to differences in regulation or the implementation of new charging principles.

REFERENCES

Bajari, P., McMillan, R., & Tadelis, S. (2009). Auctions versus negotiations in procurement: an empirical analysis. *Journal of Law, Economics, and Organization*, *25*(2), 372–399. doi: 10.1093/jleo/ewn002.

Bundesnetzagentur. (2013). Railway market analysis. *Federal Network Agency for Electricity, Gas, Telecommunications, Post and Railway*. Available online: https://bit.ly/2KYPjvG (last accessed May 4, 2020).

Casullo, L. (2017). Rail funding and financing. *Review of Network Economics*, *16*(2), 125–141. doi: 10.1515/rne-2017-0038.

Catalano, G., Daraio, C., Diana, M., Gregori, M., & Matteucci, G. (2019). Efficiency, effectiveness, and impacts assessment in the rail transport sector: a state-of-the-art critical analysis of current research. *International Transactions in Operational Research*, *26*(1), 5−40. doi: 10.1111/itor.12551.
Cowie, J. (2009). The British passenger rail privatisation: conclusions on subsidy and efficiency from the first round of franchises. *Journal of Transport Economics and Policy*, *43*(1), 85−104.
ECMT (1998). *Rail Restructuring in Europe*. Paris: OECD Publishing.
ECMT (2001). *Railway Reform: Regulation of Freight Transport Markets*. Paris: OECD Publishing.
European Commission (1991). Council Directive 91/440/EEC of July 29, 1991 on the development of the community's railways. *Official Journal of the European Union L 237, 24.08.1991*, 25–28.
European Commission (2007). Regulation (EC) No 1370/2007 of the European Parliament and of the Council of October 23, 2007 on public passenger transport services by rail and by road and repealing Council Regulations (EEC) Nos 1191/69 and 1107/70. *Official Journal of the European Union L 315, 02.12.2007*, 1−13.
European Commission (2008). Community guidelines on state aid for railway undertakings. *Communication from the Commission, Official Journal of the European Union C 184, 22.07.2008*, 13−31.
European Commission (2011). White Paper on Transport: Roadmap to a single European transport area: towards a competitive and resource-efficient transport system. *COM(2011) 144 final, Brussels, 28.3.2011*.
European Commission (2012). Directive 2012/34/EU of the European Parliament and of the Council of November 21, 2012 establishing a single European railway area. *Official Journal of the European Union L 343, 14.12.2012*, 32−77.
European Commission (2016a). Directive (EU) 2016/2370 of the European Parliament and of the Council of December 14, 2016 amending Directive 2012/34/EU as regards the opening of the market for domestic passenger transport services by rail and the governance of the railway infrastructure. *Official Journal of the European Union L 352, 23.12.2016*, 1−17.
European Commission (2016b). Regulation (EU) 2016/2338 of the European Parliament and of the Council of December 14, 2016 amending Regulation (EC) No 1370/2007 concerning the opening of the market for domestic passenger transport services by rail. *Official Journal of the European Union L 354, 23.12.2016*, 22−31.
European Commission (2018). Authorisation for state aid pursuant to Articles 107 and 108 of the Treaty on the Functioning of the European Union − cases where the Commission raises no objections. *Official Journal of the European Union C 158, 04.05.2018*, 1−9.
European Commission (2019). Authorisation for state aid pursuant to Articles 107 and 108 of the Treaty on the Functioning of the European Union − cases where the Commission raises no objections. *Official Journal of the European Union C 14, 11.01.2019*, 1−7.
European Conference of Ministers of Transport (Ed.) (2007). *Competitive tendering of rail services*. Paris: ECMT: Distributed by OECD Publishing (OCLC: ocn181345987).
IRG-rail (2017). Fifth annual market monitoring report. *Independent Regulators' Group Rail*.
IRG-rail (2018). Updated review of charging practices for the minimum access package in Europe. *Independent Regulators' Group Rail*. Version 5: November 2018.
IRG-rail (2019). Seventh annual market monitoring report. *Independent Regulators' Group Rail*.
Nash, C. (2017). Railway finance in Europe. *Review of Network Economics*, *16*(2), 67–88. doi: 10.1515/rne-2017-0039.
Nash, C., Crozet, Y., Link, H., Nilsson, J.-E., & Smith, A. (2018). Track access charges: reconciling conflicting objectives. *Project Report, Professor Chris Nash, CERRE & University of Leeds*.
Nash, C., & Wolański, M. (2010). Workshop report – benchmarking the outcome of competitive tendering. *Research in Transportation Economics*, *29*(1), 6−10. doi: 10.1016/j.retrec.2010.07 .003.
Sánchez-Borràs, M., Nash, C., Abrantes, P., & López-Pita, A. (2010). Rail access charges and the competitiveness of high speed trains. *Transport Policy*, *17*(2), 102−109. doi: 10.1016/j.tranpol.2009.12.001.
Schäfer, J.T., & Götz, G. (2017). Public budget contributions to the European rail sector. *Review of Network Economics*, *16*(2), 89–123. doi: 10.1515/rne-2017-0044.

Zunder, T.H., Islam, D.M.Z., Mortimer, P.N., & Aditjandra, P.T. (2013). How far has open access enabled the growth of cross border pan European rail freight? A case study. *Research in Transportation Business & Management*, *6*, 71–80.

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 statutes of the Justus Liebig University Giessen for ensuring good scientific practice.

Gießen, December 12, 2024

Jan Thomas Schäfer