Five Essays on Industrial Organization and Economic Policy

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades

an der Wirtschaftswissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

Heinrich Heine Universität Düsseldorf

von:	**Kai Fischer, M.Sc.**
	geboren am 15.02.2000 in Düsseldorf
Erstbetreuer/-gutachter:	Prof. Dr. Justus Haucap, DICE
Zweitbetreuer/-gutachter:	Prof. Dr. Joel Stiebale, DICE

Abgabedatum: 21. März 2025

Acknowledgements

On my journey towards a PhD in Economics (and my almost eight years at DICE), I was fortunate to receive large support from various sides. This note is my attempt to express my appreciation and gratitude.

First and foremost, I extend my thanks to my main supervisor, Justus Haucap. His encouragement fostered my passion for policy-relevant economic research since my early studies. Despite not knowing me initially, he offered me a position at DICE starting from my first semester, nurtured my interest in industrial organization, and even invited me to coauthor work during my bachelor's studies. His guidance and support kept me in Düsseldorf, even after I almost "quasi"-left after my bachelor's degree, and helped throughout the PhD. Importantly, he also arranged an enriching research stay in Leuven.

I am also grateful to Joel Stiebale, who agreed to be my second supervisor and referee for this dissertation. His advice was invaluable as I wrote my job market paper and when I navigated the job-market application process.

I owe a significant debt of gratitude to Andreas Lichter and Simon Martin. They pushed me to streamline my partly incoherent research portfolio, offered job market advice, and helped me with many academic challenges I faced.

Speaking of Simon, special thanks go to my co-authors of this thesis: Henrique Castro-Pires, Simon Martin, Marco Mello, Giuseppe Moscelli, and Philipp Schmidt-Dengler. Their collaboration was crucial to successfully completing the papers in this dissertation. I learned immensely from our joint work, especially from the research with Simon and Philipp. Additionally, Philipp's support during the job market process was indispensable, and I extend my deepest thanks to him.

I also wish to thank Frank Verboven, who generously hosted me for half a year in Leuven. This time and the experiences I collected were pivotal in the process of growing as a researcher. I am especially grateful to Frank for his efforts on my behalf during the job market season.

Thanks also go to James Reade, who invited me to Reading in the summer of 2022.

There are many more colleagues and friends I wish to thank for their time, dedication, and camaraderie. To ensure that I do not accidentally omit anyone important, I look forward to expressing my gratitude in person!

Finally, my deepest thanks go to my family - my parents and my twin brother. My parents have always been there for me, have covered my back, and have had an open ear for me all the time. Importantly, they have reminded me of the importance of life beyond academia. My brother, who shares so many of my hobbies and interests (as well as the pleasures and downsides of academia), has been a constant source of support and motivation. Having him by my side has made life (and the PhD journey as a part of it) much easier and joyful.

Thank you!

Coauthors and Publications

Chapter 1: Alcohol Prohibition and Pricing at the Pump

Coauthor(s): none
Status: This chapter has been published in the *Journal of Industrial Economics.*

Chapter 2: The Heterogeneous Effects of Entry on Prices

Coauthor(s): Simon Martin, Philipp Schmidt-Dengler (both University of Vienna)
Status: This chapter is conditionally accepted in the *Journal of the European Economic Association.* The data replication package was resubmitted to the journal and is currently being assessed by the data editor.

Chapter 3: Indirect Taxation in Consumer Search Markets: The Case of Retail Fuel

Coauthor(s): Simon Martin, Philipp Schmidt-Dengler (both University of Vienna)
Status: This chapter is available as a working paper (*CEPR Discussion Paper No. 19095*). However, the most recent revised version is not available online but is included in this dissertation.

Chapter 4: Industrial Policy in Declining Industries: Evidence from German Coal Mines

Coauthor(s): none
Status: To be made available as a discussion paper.

Chapter 5: Immigration, Workforce Composition, and Organizational Performance: The Effect of Brexit on NHS Hospital Quality

Coauthor(s): Henrique Castro-Pires (Miami Herbert Business School), Marco Mello (University of Aberdeen), Giuseppe Moscelli (University of Surrey)
Status: To be made available as a discussion paper.

Contents

Introduction

Understanding how market outcomes are influenced by external, to market participants exogenous conditions is a key objective of economics. Some of these conditions can be actively shaped by lawmakers and policymakers — for example, through taxation, regulation, or consumer policy. While many of these interventions do not prioritize competition as a primary objective (e.g., public economics policies like taxation or health policies target other certain welfare dimensions), they (in)directly impact competition and classical outcomes of industrial organization in the affected industries and markets. Given the increasing tendency to intervene in industries in recent years — often in response to growing concerns about globalization and market power — studying the diverse and complex effects of these policies is highly pertinent. This research should inform policymakers (i) ex-ante about potential implications of interventions on market structures and competition, and (ii) retrospectively help to assess already implemented policies across various dimensions.
This dissertation explores the intended and unintended consequences of economic policy interventions on competition, examining both consumer-side and firm-side behavior. In doing so, I will adopt the standard approach in industrial organization to zoom in on specific markets and industries. More specifically, I will focus on large-scale industries that are relevant to a broad portion of the population. I will study energy markets, including gasoline retailing and coal production — energy markets that are frequently subject to policy interventions. Additionally, I will investigate the healthcare supply sector and the unintended consequences of labor market policies on public health.
This thesis consists of five chapters, each addressing a specific research question but all centering around the common themes of economic policy evaluation, competition, and industrial organization. I will outline the chapters below.

The *first chapter* examines how public policy interventions in the product portfolio of firms influence their strategic behavior and pricing strategies. Specifically, I investigate the impact of product variety at gasoline stations on gasoline prices and opening hours. To date, both the academic literature and competition authorities have treated gasoline stations as single-product firms. However, gasoline stations do not solely sell (typically low-margin) fuel; they also generate significant portions of their profits from services (e.g., car wash and shop).
I focus on a nightly alcohol ban in the shops of gasoline stations in one German federal state, which was originally implemented to reduce youth alcohol consumption at night (Marcus and Siedler, 2015). By comparing opening hours and prices before and after the ban was lifted in treated and untreated federal states, I demonstrate that an active restriction on shop products also led to lower prices at the pump. This effect occurred despite weaker competition during the prohibition period, as not all gasoline stations were open at night due to fewer alcohol sales.
My analysis serves several purposes. First, it helps competition authorities to better understand the competitive environment of gasoline stations, thereby clarifying the relevant market. Second, it provides evidence that gasoline price apps, which guide consumers to cheap gasoline stations, may mislead consumers into choosing low-service, low-quality stations. Third, it informs policymakers about the unintended consequences of alcohol prohibition laws.

In the *second chapter*, my coauthors and I explore one of the central questions in industrial organization: What is the effect of competition on prices and consumer welfare? However, we extend this seemingly simple question by examining the heterogeneity in this relationship across different consumer groups, particularly those with different levels of information about product characteristics. We investigate market entry and exit in local gasoline markets, an industry often studied as a laboratory for consumers with imperfect price information (e.g., Chandra and Tappata, 2011, Martin, 2024).
We find that prices are sensitive to changes in competition, especially at the left end of the price distribution. Informed consumers, who possess better knowledge of gasoline prices, tend to purchase more frequently from the cheaper left end and thus benefit more from market entry and suffer more from exit compared to uninformed consumers.
This result highlights that average policy effects can obscure the distributional consequences across consumer groups. Moreover, the benefits of competition are more pronounced for informed consumers, which suggests that price transparency tools, such as gasoline apps, can enhance the competitive gains for marginally more informed consumers.

In the *third chapter*, I continue my research on the German gasoline industry. Together with my coauthors, we investigate how tax policy influences price setting and consumer search under conditions of imperfect price information. This question is particularly important because (i) taxes on gasoline have been adjusted at various times to alleviate consumer burden during periods of inflation, and due to (ii) the coexistence and mixture of excise and value-added taxes in this market. Additionally, the public debate surrounding gasoline taxation (in Europe) has been particularly intense during periods of rising prices, such as during the COVID-19 pandemic and following the decline in oil imports from Russia.
We provide both reduced-form and structural empirical evidence to answer this question. First, we show that a reduction in the value-added tax (VAT) on gasoline was largely passed on to consumers. However, we also find that high-income regions benefited more from this tax reduction than low-income regions, which was an unintended consequence of the policymakers' actions. We explain the heterogeneity across income groups by presenting descriptive evidence on search efforts and the potential gains from searching; both of these vary by income level.
Second, we analyze the optimal mix of excise and value-added taxes using a structural model of the industry. The model incorporates consumers' imperfect information by endogenously estimating the distribution of search costs across markets with varying levels of competition and income. Our findings suggest that, when holding tax revenues constant, cuts in excise taxes are more welfare-enhancing than the observed VAT reduction.
Also, our structural model suggests that future increases in CO_2 prices (as part of the excise tax) will be especially paid by high-income regions.

While the previous chapters focused on the gasoline retail industry as a real-world laboratory, the *fourth chapter* examines the German coal mining industry from the 1950s to 1970s. In this part of my dissertation, I address the question of how industrial policy — specifically, active policymaker interventions — can influence and shape productivity dynamics and technology adoption in declining industries. This question is highly relevant to many policymakers today, as they face challenges in industries undergoing decline, such as the shrinking car manufacturing sector in Germany or steel production in the Western world. Moreover, industrial policy has seen a resurgence in recent years (e.g., Juhasz et al., 2023).

I focus on a policy from the early 1960s that subsidized mine closures. The government offered a flat exit premium to any coal firm willing to close a mine, which resulted in the almost immediate exit of nearly one-fourth of all mines in the industry.

The analysis in this chapter demonstrates that the policy increased average, output-weighted productivity in the industry through three main channels: within-mine productivity growth, the reallocation of production to more productive mines, and the closure of unproductive mines. The within-mine productivity growth was primarily driven by increased investments, made possible by the exit subsidies that lifted the financial constraints of multi-mine firms. These efficiency gains led to significant marginal cost savings in the industry.

My findings show that industrial policy in declining industries can improve productivity and foster technology adoption — a point that has been widely debated in the literature (Rodrik, 2004). Furthermore, I demonstrate that the lifespan of an industry can be indirectly extended through policy measures without discouraging firms from making investments.

In the final *fifth chapter*, I examine the unintended consequences of migration policies on public health. In the face of demographic change, Western countries must rely on foreign healthcare workers to meet demand. However, to accurately assess the need for economic policies aimed at attracting foreign nurses, it is essential to quantify the benefits of having skilled foreign nurses in the country.

Together with my coauthors, I study this topic by analyzing a natural experiment. We investigate the reduction in the number of foreign EU-nationality nurses in English hospitals following Brexit. By comparing hospitals that were heavily reliant on EU nurses before the policy with those less dependent on them over time, we track their healthcare performance and its connection to the net outflow of EU nurses. Our findings reveal that hospital quality (measured by metrics such as fatalities, unexpected readmissions, etc.) deteriorated persistently after Brexit when a high share of nurses used to be of a non-British EU nationality in a hospital. The primary cause for the reduction in healthcare quality is the substitution of skilled EU nurses with less skilled Asian nurses.

These results show that rising protectionism tendencies (such as Brexit), next to the well-known direct effect on economic production and trade (e.g., Born et al., 2019, McGrattan and Waddle, 2020), can have substantial consequences for the economic performance and welfare of countries through unintended channels such as public health.

This thesis is a step toward a better understanding of various economic policies and their effects. However, it only covers the role of policymaking in a handful of industries and case studies. So, there is much more to explore. I am looking forward to seeing and contributing to this research in the future.

Bibliography

Born, B., Müller, G. J., Schularick, M. and Sedláček, P. (2019), 'The Costs of Economic Nationalism: Evidence from the Brexit Experiment', *The Economic Journal* **129**(623), 2722–2744.

Chandra, A. and Tappata, M. (2011), 'Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets', *The RAND Journal of Economics* **42**(4), 681–704.

Juhasz, R., Lane, N., Oehlsen, E. and Pérez, V. C. (2023), 'The Who, What, When, and How of Industrial Policy: A Text-Based Approach'. Working Paper.

Marcus, J. and Siedler, T. (2015), 'Reducing Binge Drinking? The Effect of a Ban on Late-night Off-premise Alcohol Sales on Alcohol-related Hospital Stays in Germany', *Journal of Public Economics* **123**, 55–77.

Martin, S. (2024), 'Market Transparency and Consumer Search—Evidence from the German Retail Gasoline Market', *The RAND Journal of Economics* **55**(4), 573–602.

McGrattan, E. R. and Waddle, A. (2020), 'The Impact of Brexit on Foreign Investment and Production', *American Economic Journal: Macroeconomics* **12**(1), 76–103.

Rodrik, D. (2004), 'Industrial Policy for the Twenty-first Century'. Working Paper.

Chapter 1

Alcohol Prohibition and Pricing at the Pump

Coauthor(s): none

Abstract: Firms often sell a transparent base product and a valuable add-on. If only some consumers are aware of the latter, the add-on's effect on the base product's price will be ambiguous. Cross-subsidization between products to bait uninformed consumers might lower, intrinsic utility from the add-on for informed consumers might raise the price. We study this trade-off in the gasoline market by exploiting an alcohol sales prohibition at stations as an exogenous shifter of add-on availability. Gasoline margins drop by 5% during the prohibition. The effect is mediated by shop variety and competition. Using traffic data, we unveil sizeable consumer-side reactions.

1.1 Motivation

The literature on gasoline markets is broad and has examined many features typical of gasoline competition, such as price dispersion, asymmetries in input cost pass-through and Edgeworth cycles. Most approaches to these topics assume that competition occurs only among gasoline stations, which are usually treated as single-product firms solely selling homogenous gasoline. Only a few papers have dealt with the relation of gasoline prices to stations' attached services and secondary products such as shops, supermarkets, or carwashes (Doyle et al. (2010), Haucap et al. (2017*a*,*b*), Wang (2015), Zimmerman (2012)). However, potential interactions of pricing at the pump and the provision of complementary products have relevant implications for market definition and unveil distributional consequences for heterogeneously informed consumers. If such complementarities distort the signal, that low prices imply the best deal in a homogenous product market like the gasoline market, the matching of consumers, who are uninformed about the availability of complementary products, to suitable stations could deteriorate. Also, common price transparency regulations in gasoline markets, that increase the prominence of stations with cheap gasoline prices, might be misleading then.

Whether the existence of a complementary product raises or lowers gasoline prices - relative to a world without the complement - if only some consumers are aware of the complement, is unclear from an ex-ante perspective. On the one hand, better services or a wider product assortment increase the intrinsic utility of some consumers' shopping. This can cause an outward shift in gasoline demand. Also, consumers will face opportunity costs of traveling if they are not one-stop shoppers but consume gasoline and the complement from different stations. This would explain price increases for gasoline. On the other hand, gasoline stations might use low and transparent gasoline prices as a quasi-loss leader to bait uninformed consumers, who ex-ante do not intend or expect to, in the end, buy additional products in the store. Cross-subsidization could arise (Armstrong and Vickers (2012), Gabaix and Laibson (2006), Heidhues et al. (2017), Lal and Matutes (1994)). Less transparently priced complementary products such as add-on services or shop products might then be purchased by consumers at relatively high prices. Therefore, the overall price effect of complementary products on gasoline prices is ambiguous and a question for empirical research. Similar trade-offs can be found in most markets.

In this work, we go into this matter and answer the question of how the introduction of a complementary product affects a firm's price setting for other products. We provide causal evidence by exploiting a unique setting in the gasoline market, where the availability of a complement is exogenously determined by public policy. In particular, we examine a quasi-experiment, the lifting of a local nightly alcohol ban at gasoline stations in a federal state of Germany, as a shifter of complement availability. The prohibition restricted the shop assortment of stations as it mandated sales of alcohol, an important add-on product for gasoline stations, to be forbidden from 10pm to 5am. The policy was implemented in 2010 and lifted in December 2017. It aimed at the reduction of binge alcohol consumption among youths at night. As 60% of all profits of German gasoline stations are linked to the shop, 20% to carwashes, and only 20% to gasoline sales (FAZ (2015), Ivanov (2019), Nicolai (2021), NTV (2015)), the alcohol sales ban reflects a relevant revenue shock.

To analyse the effect of the available add-on on the price of the complementary base product gasoline, we use real-time data of all gasoline prices in the German gasoline market at the station level. By means of a difference-in-differences setup, we take advantage of the low menu costs and within-day variation of prices and compare gasoline prices during and after the prohibition as well as between affected and unaffected stations. This allows us to unveil the overall price effect of add-on availability and, hence, the complementarity on the base product's price. Building on precise information about stations' competitive environment and brand affiliation, we further can investigate heterogeneity across firms.

Our findings and contributions to the literature are threefold. Firstly, we investigate the effect direction of add-on quality on gasoline prices. We find nightly prices of stations affected by the prohibition to increase by 0.6 Eurocent/l - or 5% of the gross margin - after the lifting of the prohibition. Hence, especially consumers who did not buy alcohol profited from the policy when it was in place. Stations with smaller product variety, where alcohol's relative importance for shop revenues is higher, reveal even stronger price effects. Similarly, stations with few competitors nearby increase prices more strongly. Opportunity costs of buying alcohol at another station increase with decreasing competition intensity. Thus, a potential cross-subsidization mechanism is overall outweighed by the intrinsic value of additional services. Using detailed, geo-coded traffic counter data, we provide supporting evidence that traffic increases only in the direct vicinity of gasoline stations after the reintroduction of alcohol sales.

Our findings add to the literature on the role of station amenities for stations' pricing behavior. Other papers have shown that stations' choice to operate convenience stores (Doyle et al. (2010), Ning and Haining (2003), Haucap et al. (2017*a*)) and the proximity to hypermarkets nearby (Zimmerman (2012)) indeed shape pricing behavior. Though, they mainly rely on the endogenous self-selection of stations into low- or high-quality segments while we exploit an exogenous shifter of service and add-on availability. Our results also address the delineation of gasoline markets as price effects vary with the exposure to alcohol sales. Alcohol revenues are also determined by local supermarkets or pubs. This indicates that gasoline stations might not only compete with other stations.

Secondly, while our results address discussions on multi-product competition across most markets, note that the setting studied in this paper is unique. It mainly differs from other markets with two price components in three ways: At first, add-on services often are valueless to the consumer and are only jointly bought with the base good such as overdraft fees for financial services (Armstrong and Vickers (2012), Gabaix and Laibson (2006)). In our setting, consumers are free to opt out of buying alcohol but can still buy other shop products. Beyond that, purchasing alcohol gives positive utility to some consumers. Second, firms often endogenously set the prevailing level of consumer information about prices in the market for the base product by, for example, advertising prices. We consider a price transparency environment that exogenously dictates prices to be equally transparent across firms. By law,

gasoline prices of all German stations are published in real-time for consumers. Lastly, we do not just vary add-on revenues but study the add-on existence at the extensive margin. Hence, our results represent an upper bound for fluctuations of add-on revenues in our setting and are helpful in forming benchmarks for other industries.
Thirdly, we analyze how active stations are in response to the prohibition. 10% more stations adjust prices during night hours after the prohibition lifting. While this observation could purely represent changes in the Edgeworth cycles, we show that prominent characteristics of price cycles are unaffected by the lifting. Therefore, we believe these findings express changes in opening hours.
The remainder of the paper is as follows: We start with an explanation of the institutional background and a theoretical motivation in Sections 1.2 and 1.3 before presenting our data and empirical strategy in Section 1.4. We then proceed with our analysis in Section 1.5 before providing robustness checks and a conclusion in Sections 1.6 and 1.7.

1.2 Institutional Background

Particularly, we examine a nightly off-premise alcohol prohibition in Baden-Wuerttemberg, a German federal state with a population of eleven million. This policy primarily affected gasoline stations as the main nightly off-premise places to go for alcohol (Marcus and Siedler (2015), Baueml et al. (2023))[1]. From 2010 onwards, Baden-Wuerttemberg prohibited nightly alcohol sales from 10pm to 5am via the 'Alkoholverkaufsverbotsgesetz' (Alcohol Sales Prohibition Law). As most people do not prestore alcohol, the prohibition was binding (Marcus and Siedler (2015)). This specific legislation ran out on December 08, 2017, as local authorities from then on should have selected specific 'hotspots' (e.g., city centres) for bans only. In the three years after the lifting of the policy, there, though, were only rare occasions, when a municipality implemented an alcohol consumption prohibition - mainly during festivals (Landtag von Baden-Wuerttemberg (2020)).
Its main intentions were the reduction of binge drinking among youths and of indirect spillovers on crime (Baumann et al. (2019), Baueml et al. (2023), Marcus and Siedler (2015)). The policy was effective in several ways indicating a real shock in the volume of alcohol consumed. Up to now, Baueml et al. (2023), Marcus and Siedler (2015) and Baumann et al. (2019) discussed direct effects on health costs (hospital admissions, doctor visits) and crime for this specific case study. All three papers find that the policy had an economically relevant effect. The number and the length of hospital stays among youth binge drinkers and late-night assaults fell due to the policy. The effect is strongest on young adults since they are more price-sensitive, can hardly pre-store alcohol in their parents' home and are more likely to conduct off-premise pre-drinking (Baueml et al. (2023)).
As the legislation ran out ahead of time - it was expected that the legislation would not

[1]During the prohibition, only stations that also ran a diner with an official catering license to sell on-premise alcohol were still allowed to sell alcohol at night (§3a Abs. 1 LadÖG). This mainly concerned highway stations with rest houses, which at the same time were not allowed to sell alcohol due to a highway-specific alcohol prohibition.

change before 2018 (Mayer (2017)) - and because the law was ineffective just a few days after the public announcement of the abolition, anticipatory effects are unlikely.

We expect such regulation to have a sizeable impact on the German gasoline market. In Europe, German gasoline stations have one of the lowest net margins on fuels (Scope Ratings (2019)). Therefore, shop sales make up a relevant share of stations' overall profits. In particular, alcohol and beverage sales account for more than 10% of all in-shop sales (Scope Ratings (2019)). Moreover, consumers coming for alcohol buy other products on the way. Recent years have shown that especially big brands such as ARAL extended their shops by for example integrating shops of supermarket chains. In contrast to other countries, German gasoline stations mostly did not introduce paying at the pump by card, as this would stop consumers from entering the store. Hence, most stations are occupied in person all day long, so that shop sales are possible. Moreover, German gasoline stations often act as "shopping location of last resort" during night times as then German groceries rarely open. Thus, a nightly prohibition impedes a relevant business time.

Alcohol revenues may be relevant for gasoline prices. In response, cross-subsidization could plausibly be an optimal pricing strategy next to quality-related price inclines. To show this, we perform a simple, hypothetical back-of-the-envelope calculation based on some assumptions. Following the Statistisches Landesamt Baden-Württemberg (2022), overall annual gasoline and diesel consumption was approximately 7 million tonnes or 9 billion litres in 2017. Admittedly gasoline demand is low at night. But the Federal Cartel Office (2019) documents that still around 5% of all car drivers preferably fuel at night (10pm to 5am), which gives a lower bound of the actual demand. This implies that at least around 425 million litres p.a. are sold in Baden-Wuerttemberg at night. Uniformly distributing this over approximately 800 gasoline stations which operate at this daytime, this is slightly more than 0.5 million litres per station and year. If a station followed a cross-subsidization strategy that lowers margins by, for example, only half a Eurocent/l, it would lose around 2,500 Euro p.a.. This needs to be compensated by additional alcohol sales triggered through lower prices at the pump. Following Scope Ratings (2018), German gasoline stations, on average, earn almost one million Euro shop revenues p.a., of which alcohol products account for approximately a tenth. As alcohol is sold in the evening and night hours for the most part, profits from alcohol sales due to additional attracted consumers could exceed the cost of using gasoline as bait. In the setting studied in this paper, consumers' alcohol demand response to lower gasoline prices is changed from zero to potentially non-zero after lifting the prohibition.

1.3 Theoretical Sketch

To get a better understanding of the ex-ante ambiguity of the policy's effect on gasoline prices, we consider the differences between a gasoline station's optimization problem before and after the policy lifting. Before the lifting, the station can only sell gasoline. After the lifting, alcohol can be sold in addition.

We model a market in which consumers have heterogeneous preferences for alcohol and differ in whether they anticipate buying alcohol at a gasoline station or not. The model is set up in the following way: on the consumer side, a share of α consumers only want to buy gasoline and no alcohol. λ is the share of informed consumers - among those who do potentially buy alcohol - who are aware of the availability of alcohol products when choosing a gasoline station. $(1-\alpha)(1-\lambda)$ consumers do not consider the existence of alcohol at all. The demand of only-gasoline consumers is given by $D_\alpha(p_G^t)$ with p_G^t being the gasoline price before $(t=b)$ and after $(t=a)$ the lifting. For consumers who potentially buy alcohol, informed and uninformed consumers' demand is given by $D_{1-\alpha,\lambda}(p_G^t,\gamma_A)$ and $D_{1-\alpha,1-\lambda}(p_G^t)$ with $t \in \{a,b\}$ respectively. $\gamma_A \in \{0,1\}$ indicates whether alcohol is available $(\gamma_A = 1)$ or not $(\gamma_A = 0)$. Alcohol can only be sold after the lifting. If consumers gain utility from alcohol, then $D_{1-\alpha,\lambda}(\hat{p},1) > D_{1-\alpha,\lambda}(\hat{p},0)\ \forall\ \hat{p}$, i.e. alcohol availability causes an outward shift in gasoline demand. For simplicity, we assume (marginal) costs of zero.[2]

We then construct the gasoline station's profit function before (π^b)

$$\pi^b = p_G^b[\underbrace{\alpha D_\alpha(p_G^b)}_{(1)} + \underbrace{(1-\alpha)(\lambda D_{1-\alpha,\lambda}(p_G^b,0)}_{(2)} + \underbrace{(1-\lambda)D_{1-\alpha,1-\lambda}(p_G^b))}_{(3)}] \tag{1.1}$$

and after the prohibition lifting (π^a)

$$\pi^a = (p_G^a + p_A)[(1-\alpha)(\lambda D_{1-\alpha,\lambda}(p_G^a,1) + (1-\lambda)D_{1-\alpha,1-\lambda}(p_G^a))] + p_G^a \alpha D_\alpha(p_G^a). \tag{1.2}$$

Before the prohibition lifting, the station earns the price p_G^b from (1) those consumers who are only willing to buy gasoline and from (2) informed and (3) uninformed consumers who would also buy alcohol if available. After the prohibition lifting, the gasoline station is paid p_G^a by the same three groups. In addition, they earn alcohol revenues from those informed and uninformed willing to buy it. Also, the station faces an outward shift in the demand for gasoline from informed consumers due to the add-on availability.

Maximizing profits and rearranging the first-order conditions yields the policy's price effect.

Result 1. The price effect of the policy, Δp_G, is implicitly given by the expression

$$\Delta p_G = p_G^a - p_G^b = \Delta_{CS} + \Delta_{SQ} \tag{1.3}$$

$$\text{where } \Delta_{CS} = -\frac{p_A[(1-\alpha)(\lambda\frac{\partial D_{1-\alpha,\lambda}(p_G^a,1)}{\partial p_G^a} + (1-\lambda)\frac{\partial D_{1-\alpha,1-\lambda}(p_G^a)}{\partial p_G^a})]}{\alpha\frac{\partial D_\alpha(p_G^a)}{\partial p_G^a} + (1-\alpha)(\lambda\frac{\partial D_{1-\alpha,\lambda}(p_G^a,1)}{\partial p_G^a} + (1-\lambda)\frac{\partial D_{1-\alpha,1-\lambda}(p_G^a)}{\partial p_G^a})}$$

$$\text{and } \Delta_{SQ} = \frac{\alpha D_\alpha(p_G^b) + (1-\alpha)(\lambda D_{1-\alpha,\lambda}(p_G^b,0) + (1-\lambda)D_{1-\alpha,1-\lambda}(p_G^b))}{\alpha\frac{\partial D_\alpha(p_G^b)}{\partial p_G^a} + (1-\alpha)(\lambda\frac{\partial D_{1-\alpha,\lambda}(p_G^b,0)}{\partial p_G^b} + (1-\lambda)\frac{\partial D_{1-\alpha,1-\lambda}(p_G^b)}{\partial p_G^b})}$$
$$- \frac{\alpha D_\alpha(p_G^a) + (1-\alpha)(\lambda D_{1-\alpha,\lambda}(p_G^a,1) + (1-\lambda)D_{1-\alpha,1-\lambda}(p_G^a))}{\alpha\frac{\partial D_\alpha(p_G^a)}{\partial p_G^a} + (1-\alpha)(\lambda\frac{\partial D_{1-\alpha,\lambda}(p_G^a,1)}{\partial p_G^a} + (1-\lambda)\frac{\partial D_{1-\alpha,1-\lambda}(p_G^a)}{\partial p_G^a})}.$$

[2] We also ignore the add-on's price p_A in $D_\lambda(p_G^t,\gamma_A)$ which does not affect the sign of the prohibition's price effect as long as the intrinsic utility from the add-on is sufficiently high.

The expressions Δ_{CS} and Δ_{SQ} represent the two channels that mainly drive price differences for gasoline before and after the policy lifting: Firstly, prices after the lifting are reduced as stations cross-subsidize between alcohol and gasoline revenues. This is expressed in the first addend of (1.3), Δ_{CS}, which is negative. This term expresses that per-consumer alcohol revenues p_A are negatively correlated with the gasoline price p_G^a. This *cross-subsidization channel* characterizes gasoline as bait for uninformed consumers. Secondly, informed consumers increase demand due to the availability of alcohol products. This is expressed in the difference between the two addends of Δ_{SQ}, where alcohol availability ($\gamma_A = 1$) increases demand after the policy change (s. nominator). This *service quality channel*, hence, increases demand and prices. Thus, the overall effect on Δp_G is ambiguous.
The model further delivers intuitive predictions on how different parameters mitigate the size of the price effect or determine its sign:

Result 2. The treatment effect Δp_G

- increases in the alcohol-induced demand shift of informed consumers $D_\lambda(p_G^a, 1) - D_\lambda(p_G^a, 0)$,
- is (weakly) negative in perfectly uninformed markets ($\lambda = 0$) and
- vanishes in markets with only gasoline buyers: $\lim_{\alpha \to 1} \Delta p_G = 0$.

Result 2 states the following: If more consumers are aware of the utility gain from the availability of alcohol, this strengthens the demand expansion of the *service quality channel*. This is the case when $D_\lambda(p_G^a, 1) - D_\lambda(p_G^a, 0)$ increases. The channel will be non-existent if no consumer is aware of alcohol ($\lambda = 0$). However, the *cross-subsidization channel* is fostered by higher per-consumer alcohol revenues p_A in equilibrium, which can, for example, arise from an outward shift in alcohol demand. Both channels will become irrelevant if consumers only buy gasoline ($\alpha = 1$). We use these predictions to guide our empirical analysis of treatment effect heterogeneity across different types of markets and stations later on. This allows us to better understand whether observed prices support the modeled channels.
Nevertheless, the observed price effects of the policy might not be purely related to the channels discussed above. For example, we assumed that the share of informed consumers λ among potential alcohol consumers and the share of only-gasoline consumers α do not change with the policy lifting. Changes in these variables could rationalize positive as well as negative price effects of the policy lifting beyond the channels we discussed above. A higher share of informed consumers, who want to buy alcohol and gasoline, arrive for alcohol and could be less elastic with respect to the gasoline price. A change in the demand elasticity could explain price changes then. We discuss such other potential mechanisms in our empirical analysis later on to clarify the role of the channels modeled above.

1.4 Data and Empirical Strategy

Gasoline Price Data. We make use of E5 gasoline prices from all German gasoline stations. The data is collected by the Market Transparency Unit for Fuels (MTU) at the German Federal Cartel Office and accessed via `tankerkoenig.de`. The data is gathered in real time which allows us to exploit within-day price variation as needed in our setup. We use a full year of price data (mid-September 2017 to mid-September 2018). We construct the time-weighted average daytime (05am to 10pm) and nighttime (10pm to 05am) price per week and station.

Station Characteristics. Further, the MTU provides exact information on station characteristics such as their brand affiliation and geographical location. From this source, we construct several variables that, later on, guide our heterogeneity analysis. First, we derive whether stations open all day long (24/7) and operate at night which is reported in the MTU data[3]. Our final sample only consists of such 24/7 stations as other stations do not operate all night, which is the prohibition period.
Second, we use the location data to match stations to municipalities and counties. This allows us to match detailed information on municipality- and county-level variables such as population density, degree of urbanity, or the share of youths in the overall population.
Third, based on stations' brand affiliation, we identify stations' degree of upstream integration and station's brand value. We follow Federal Cartel Office (2011) in classifying stations into oligopolistic and non-oligopolistic stations as well as premium and non-premium stations. Previous research found that oligopolistic and premium stations tend to be expensive Haucap et al. (2017*a*). Using these classifications, we can proxy market power and heterogeneity in shop assortments.
Fourth, we also construct competition measures such as the distance to the nearest competitor or the number of stations in a certain radius around a station. We differentiate between daytime and nighttime competition measures. Daytime competition includes all stations nearby while nighttime competition is restricted to 24/7 stations as competitors.
We discuss most of the named variables in the descriptive statistics later on.
Finally, we manually identify around 380 highway stations from our sample as those are typically assigned to a separate market (Federal Cartel Office (2011))[4]. They also face §15 Abs. 4 Bundesfernstrassengesetz (FStrG), which prohibits selling alcohol at highway stations from 12pm to 7am, independently of the discussed prohibition. Hence, the lifting of the treatment should not have been binding as they are still not allowed to sell alcohol.
Overall, we end up with a panel of more than half a million observations for over 6,000 24/7 stations of which approximately 13% are located in Baden-Wuerttemberg.

Traffic Counter Data. We, moreover, use novel hourly traffic flow information from around 1,700 traffic counters in Germany. This data is publically available from the 'Bundesanstalt für Straßenwesen' and allows us to study the reaction of traffic flows in response to the policy. In detail, the data reports the number of cars passing by a certain counter within a specific hour for each day. We also know on which type of road the counters are located. Each counter's location is geo-coded and hence we can calculate the distance to the nearest open station or the federal state's border. While this data does not exactly reflect demand data, it can unveil traffic reactions to the policy and, by that, might help to understand the mechanism behind our findings.

[3]We extract this information from the first fully covered opening hours by the MTU being publically available from January 2019, just three months after the end of our sample period.

[4]For details on German highway stations see Haucap et al. (2017*a*) and Korff (2021).

Empirical Approach. Using this data, we apply a triple difference-in-differences (TDID) estimator, which studies the effect of abolishing the prohibition across federal states and daytimes[5]. We prefer a TDID estimator over a DID estimator with just nightly prices before and after the lifting because prices are correlated within the day due to intra-day Edgeworth cycles. So, we avoid missing treatment effects pushed out of the nighttime period (e.g. anticipatory alcohol purchases right before 10pm might affect prices). Nevertheless, we provide supporting simple DID results on daytime and nighttime prices separately later on as well. The regression setup is as follows:

$$P^{E5}_{swn} = \alpha_s + \lambda_w + \lambda_w \times Night_n + \beta_1(BW_s \times Night_n) + \beta_2(BW_s \times Post_w)$$

$$+\beta_3(BW_s \times Night_n \times Post_w) + \epsilon_{swn}$$

In particular, P^{E5}_{swn} is the E5 gasoline price at station s in week w at daytime $n \in \{Day, Night\}$. α_s and λ_w are station and week fixed effects. $\lambda_w \times Night_n$ are week-times-daytime fixed effects that control for underlying daytime and week trends[6]. $BW_s \times Night_n$ and $BW_s \times Post_w$ control for daytime and real-time price differences between control and treatment group where BW_s, $Night_n$ and $Post_w$ are dummies for (i) the treated federal state, (ii) night hours, and (iii) weeks after the date of the prohibition lifting, 08th of December 2017. $BW_s \times Night_n \times Post_w$ is the treatment indicator, so that β_3 gives the treatment effect of the policy lifting. We later on show that our results are robust to other specifications of the TDID setup. The identical regression approach will be used to study traffic flows later on.

Note that we assign the period after the policy lifting as the treatment period. While the prohibition (pre-lifting) period might also be seen as treatment, we understand the treatment to be the regained availability of alcohol sales at gasoline stations.

Identification. We observe an exogenous policy treatment on the state level. Interpreting our estimates as causal is valid under the assumptions that (i) treated and untreated stations would have been on the same trend in the absence of the treatment and that (ii) treatment and firm behavior of one station does not affect the treatment and outcomes of other stations (corresponding to the stable unit treatment value assumption). To investigate the parallel trend assumption, we will provide dynamic TDID regressions where to split up the treatment effect into its time-specific components. Flat pre-trends will be indicative of whether the parallel trend assumption is fulfilled in our setting (s. e.g., Olden and Moen (2022)). The setup is as follows:

$$P^{E5}_{swn} = \alpha_s + \lambda_w + \lambda_w \times Night_n + \beta_1(BW_s \times Night_n) + \beta_2(BW_s \times Post_w)$$
$$+ \sum_{t=\underline{\tau}, t\neq -1,-2}^{\bar{\tau}} \gamma_t 1[BW_s \times Night_n \times Lifting_{w-t}] + \epsilon_{swn}$$

[5] As the MTU has been launched after the prohibition's introduction in 2010, we study its lifting.

[6] Due to changes in the Edgeworth cycles over time, daytime price effects differ strongly in real-time, so that we control for this variation by interacting the daytime dummy with week fixed effects.

where $\sum_{t=\underline{\tau},t\neq -1,-2}^{\bar{\tau}} \gamma_t 1[BW_s \times Night_n \times Lifting_{w-t}]$ gives the sum of all leads and lags of the treatment effect - in two-week bins - except for the omitted reference category before the shock.

Regarding the second assumption, spillovers between treated and untreated stations are unlikely due to the exogenously fixed treatment, strict geographical separation of treated and untreated stations, and narrow local markets. There are only interactions between treated and untreated stations at the state border, which we will investigate later on.

Besides that, one concern in our setting is that the composition of treatment and control group changes due to the treatment. For example, fewer revenues due to the prohibition may lead to market exit during nighttime. Note that this would only downward bias our effect due to softening competition and higher prices during the prohibition[7]. If our treatment effect is positive, we, therefore, do not face problems interpreting results about which channel outweighs in the discussed trade-off.

Descriptive Statistics. As treatment effects might be a function of, for example, station characteristics or local competition, Table 1.1 offers insights into structural differences between the treatment and control group before the treatment. While the price level prior to the prohibition lifting has not been statistically different across both groups, the likelihood to operate at night in terms of changing prices was more extensive outside of Baden-Wuerttemberg. Competitive environments, on average, are similar and stations mostly seem to differ in the likelihood of being affiliated with an oligopolistic brand. These stations are meant to have high market power in, for example, steering the Edgeworth cycles (Federal Cartel Office (2011)). Lastly, treated stations are more likely to be located in wealthier counties with more vehicles per person. As station differences, hence, primarily lie in mostly time-invariant dimensions such as brand affiliation or county-specific demand conditions, we are able to address this heterogeneity by, for example, using station fixed effects.

Descriptives on traffic flow data are reported in Table 1.A.1 in the appendix. There are 132 counters in Baden-Wuerttemberg and 1,554 in other federal states. A median traffic counter is around 3km away from the nearest station, which opens 24/7 and counts around 25,000 (2,500) cars during daytime (nighttime) per day. This includes traffic on both sides of the road.

1.5 Results

Baseline Results. We present our baseline results in Table 1.2. A positive treatment effect would imply prices increase after a prohibition lifting. In this case, the direct quality-price complementarity would outperform the cross-subsidization channel. Our baseline results in model (1) show that, generally, prices rise in Baden-Wuerttemberg after the prohibition lifting during night hours. The effect size is 0.56 Eurocents/l. To put this into context, we

[7] We later on provide a discussion on the size of the potential downward bias when discussing nighttime market entry of stations in response to the policy lifting.

Table 1.1: Descriptive Statistics

Statistic	Units	Control (Pre-Lifting) (1)	BW (2)	Δ (p-value) (3)
Outcomes				
E5 Gasoline Price (Day)	Euro/l	1.373	1.370	0.39
E5 Gasoline Price (Night)	Euro/l	1.439	1.433	0.11
Margin (Day)	Euro/l	0.094	0.092	0.38
Margin (Night)	Euro/l	0.149	0.145	0.11
1[Active between 10pm and 5 am]	yes/no	0.863	0.836	0.03**
1[Active between 11pm and 4 am]	yes/no	0.515	0.472	0.01***
Competition				
# Competitors 0.5km Radius (Day)	#	0.472	0.449	0.56
# Competitors 0.5km Radius (Night)	#	0.257	0.230	0.32
# Competitors 1km Radius (Day)	#	1.081	1.070	0.90
# Competitors 1km Radius (Night)	#	0.546	0.535	0.85
Station Characteristics				
Share of Youths (18-25 y.o., County Level)		0.086	0.096	0.00***
Share of Youths (18-25 y.o., Municipality Level)		0.075	0.082	0.00***
Premium Station	yes/no	0.437	0.411	0.32
Oligopolistic Station	yes/no	0.372	0.273	0.00***
Highway Station	yes/no	0.051	0.046	0.65

Note: This table compares descriptive statistics of untreated stations with treated stations (both pre-treatment). The p-values come from linear regressions of the respective outcome on an intercept and a dummy for Baden-Wuerttemberg where we implement standard errors clustered at the county level.

calculate *gross margins* of gasoline stations[8]. We show that gross margins increase by around 5%. Note that gross margins still include transportation or variable labour costs, so that *net* margins should be affected even more strongly. Net margins mostly do not exceed two Eurocents/l (Scope Ratings (2019)).

To show that we do not take up unrelated variation, which does not correspond to the daytime-specific treatment, we check whether night prices purely drive the effect in models (2) and (3). The respective simple DID regressions show that only night prices increase significantly while day prices are unaffected by the prohibition lifting. This is in line with our intuition. In model (4), we use gross margins as an outcome, which are subject to subtracting, for example, labour costs to arrive at net margins.

A positive effect is indicative of alcohol assortments improving the quality of gasoline stations for the consumers. Consumers are willing to pay more at the pump as they, for example,

[8]We calculate stations' gross margins based on average, daily input costs data. For this, we obtain wholesale prices from the Oil Market Report by Argus Media - a source also used in Assad et al. (2021) and Haucap et al. (2017*a*). Cost data already includes the energy tax. Margins then are given by the VAT-deducted price minus the input costs. The average margin in the sample is 10.6 Eurocent/l which fits survey evidence (Scope Ratings (2018)).

Table 1.2: (Triple) Difference-in-Differences Regression

	Gasoline Price in Euro/l			ln(Gross Margin)
	(1)	(2)	(3)	(4)
BW×Night×Post	0.0056**			0.0481**
	(0.0022)			(0.0191)
BW×Post		0.0080***	0.0025	
		(0.0024)	(0.0019)	
Approach	TDID	DID	DID	TDID
Sample	Baseline	Only Night	Only Day	Baseline
Observations	593,193	296,598	296,595	593,076
Adjusted R^2	0.889	0.876	0.953	0.784

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows the regression equation from the 'Data and Empirical Strategy' section. Simple DIDs in models (2) and (3) include station and week fixed effects as well as the reported interaction term. 0.01% of all observations have a negative margin which we drop in the regression of logged margins in column (4).

get additional services. If consumers enter the station to purchase alcohol, gasoline might be sold as a by-product. Interestingly, we do not find any evidence for lower gasoline prices after the prohibition which fits a story of gasoline being a bait product for stations. This would have been in line with cross-subsidization if consumers had not been aware of buying alcohol when approaching a station to fuel (Gabaix and Laibson (2006), Lal and Matutes (1994)). Similarly, Haucap et al. (2017*b*) discussed that carwashes or supermarkets typically offer fuel cheaply. Hence, the mechanism underlying our observations here is likely to be reversed. Consumers approach stations with the purpose of buying alcohol and then are willing to fuel at a higher price as they otherwise would face non-negligible opportunity costs of an additional trip.

The effect is remarkable, especially when considering that alcohol sales only make up 10% of an average station's shop revenues. Extrapolating this to the overall importance of the shop for price setting, gasoline competition is highly related to shop revenues. Strategic interactions between shop assortment and gasoline prices also indicate that gasoline stations act like multi-product firms.

Note that we cannot fully exclude that our reduced-form effect is a sum of a cross-subsidization effect (which reduces gasoline prices) and the discussed quality improvement (which increases prices). We can only ensure that the quality and intrinsic utility channel dominates. We later check whether cross-subsidization may play out more strongly for bigger shops, so that the treatment effect might vary across stations' types of shops.

Dynamic Estimates. To verify that the observed effects really originate from the legalization's lifting and hence can be interpreted as causal, we provide two types of dynamic approaches: Firstly, we apply a dynamic TDID setup in which the treatment effect is split up into several smaller time intervals before and after the treatment.

Figure 1.1 gives the dynamic estimates from the baseline regression above. As evident, we observe that the significant price drop arises just after lifting the prohibition. While there is a slight delay until the treatment effect evolves, the effect size remains constant after some weeks until the end of the time window[9]. The slight delay is in line with the unexpected timing of the policy lifting. Pre-trends are flat which gives us certainty that the effect is a consequence of the policy change.

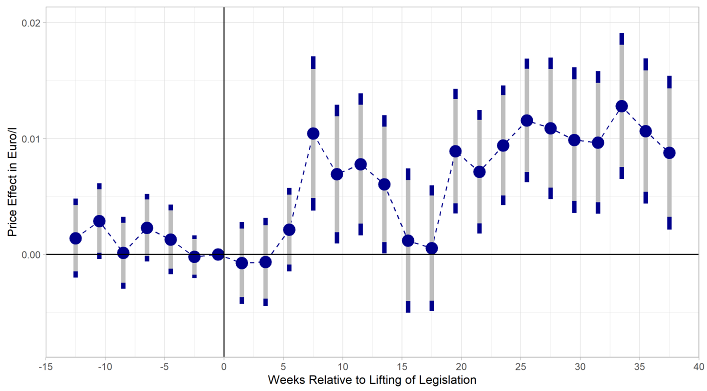


Figure 1.1: Dynamic Difference-in-Differences Estimates

Note: This plot gives dynamic estimates of γ_t from the dynamic DID strategy discussed in the 'Data and Empirical Strategy' section. The exact timing of the end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals. Standard errors are clustered at the county level.

Besides showing that the effect only arises after the legislation, we provide evidence beyond models (2) and (3) in Table 1.2 that the treatment effect is purely bounded to night hours. This is done in Figure 1.2, where we run the simple DID regression of whether prices changed in Baden-Wuerttemberg after the treatment for hourly average prices on the week level separately. Indeed, the results closely represent the hypothesis that there is no treatment effect over daytime while a treatment effect arises at night. The effect does not appear immediately after 10 pm, which is likely related to limited demand effects for alcohol. Some supermarkets are still open until midnight, and most restaurants have not closed yet. Admittedly, there is a significant effect remaining between 5 and 6am. This is likely related to the given timing of the Germany-wide intra-day Edgeworth cycles, where most stations changed prices after 6am (Federal Cartel Office (2018)).

Heterogeneity Analyses. To understand which stations are more prone to react to the prohibition, we study effect heterogeneity across station characteristics such as competition at the pump, variety in the product assortment, or brand affiliation.

Firstly, we study competition effects. As described above, our price effect likely originates from the mechanism that alcohol-demanding consumers visit gasoline stations and consume

[9]After about 15 weeks, there is a short-term drop in the effect size. The timing corresponds to the Easter holidays and hence might reflect a short-term heterogeneous exposure to demand for alcohol at gasoline stations across federal stations in Germany.

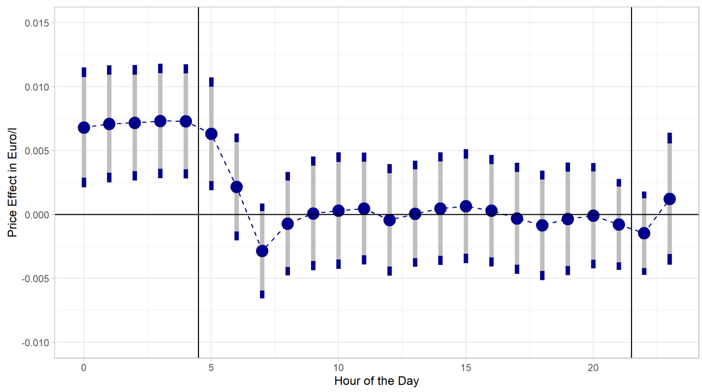


Figure 1.2: Dynamic Effects by Hour of Day

Note: This plot gives dynamic estimates of the interaction term $BW_s \times Post_w$ of a simple DID model where one regression is run for each hour separately. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. Standard errors are clustered at the county level. We provide 90 and 95% confidence intervals for all coefficients.

gasoline on the side. Then, the price effect would arise from the opportunity costs of traveling to a different gasoline outlet. This effect should be larger if alternative stations are far away. Similarly, if consumers only have one station nearby, they are more likely to be informed about the add-on which reduces the cross-subsidization incentive. Hence, lower gasoline competition should foster the effect. We study this by splitting the sample at the median number of nightly competitors in a 1km radius[10]. Figure 1.3 reports our results on heterogeneity analyses. Indeed, in the first panel of Figure 1.3, we find that lower competition is related to a higher nightly price increase after the prohibition lifting. Simultaneously, higher competition is correlated with stations lying in densely populated areas, so that stations in cities do not drive our effect[11]. In cities, alcohol consumers may be motorized less often which does not incentivize changes in gasoline prices.

Secondly, we study how ex-ante shop assortments impact the price effect's size. To sort stations into different shop categories, we follow the definition by the Federal Cartel Office (2011). Stations are sorted into premium and small assortment stations based on their brand affiliation. Premium stations are known for a wider assortment of products. Alcohol is a very simple product offered by any station, so that the marginal return and relative importance of alcohol revenues is typically higher in smaller shops. At the same time, larger shops imply larger per-consumer revenues from alcohol visitors, which fosters the cross-subsidization effect. Both arguments propose larger shops experience a lower price effect of the policy. In the second panel of Figure 1.3, we find premium stations with large product variety do not react significantly, while the price effect is especially evident for low assortment stations. Consumers who buy alcohol at gasoline stations may be likely to buy other shop products there as well, so that bigger shops do not experience a comparable shock to shops with smaller product varieties. In contrast, the premium station may face consumers who buy more af-

[10]Our results also hold for different radii and sample splits not at the median.

[11]This also holds when studying the effect heterogeneity across county differences in the population density.

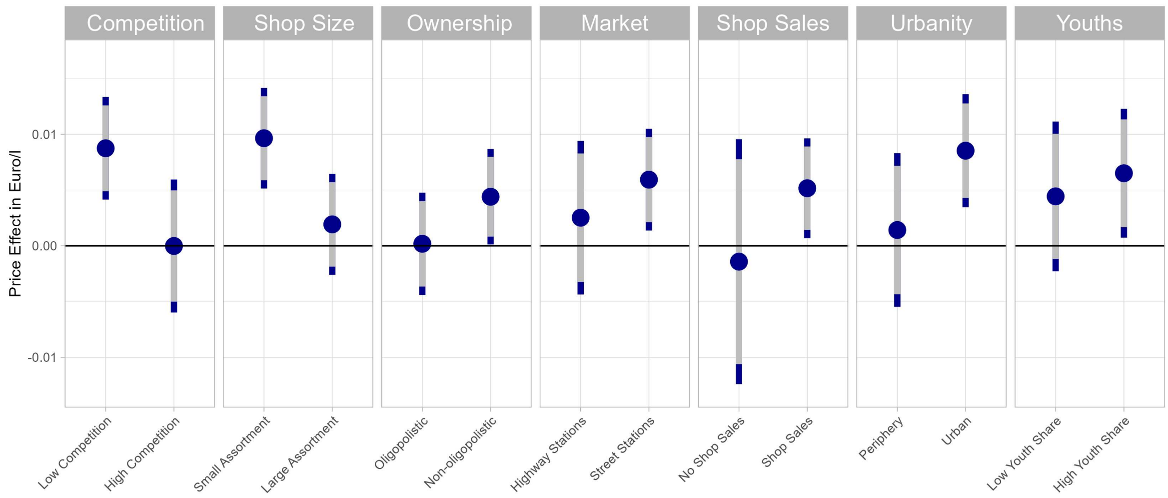


Figure 1.3: Heterogeneity Analyses: Intensive Margin

Note: This plot gives the treatment effect β_3 from the baseline regression for subsamples along firm characteristics. The y-axis documents the effect size in Eurocent/l, the x-axis gives the respective subsample. 90% and 95% confidence bands are reported. Standard errors are clustered at the county level.

ter the prohibition but have visited the station before as well. This then does not lead to more gasoline sold at premium stations. Also, the null effect for premium stations might be a result of stronger cross-subsidization since gasoline purchasers might buy more products beyond alcohol when entering the store.

Thirdly, we study the role of market power. In the German gasoline market, market power is associated with vertical integration to oil refinery firms as these also supply competitors and have been determining the daily Edgeworth cycles for years (Federal Cartel Office (2011), Siekmann (2017)). Vertically integrated, so-called 'oligopolistic' brands are, for example, Shell, Aral (BP), or Total. We study whether the effect differs across oligopolistic and non-oligopolistic brands. We find that especially non-oligopolistic brands increase nightly prices after the prohibition lifting. Our results in the third panel of Figure 1.3 show that oligopolistic stations' price level was not lower before the prohibition lifting, so that a price drop during the prohibition did not occur at stations with market power.

Fourthly, we study a sample of only highway stations in the fourth panel of Figure 1.3. Highway stations have been subject to an alcohol prohibition throughout night hours, independent of the discussed alcohol prohibition. Hence, as these stations were still not subject to the opportunity to sell alcohol from December 08, 2017, onwards, we expect to observe a zero treatment effect. In terms of our model, both channels are switched off. That is why this analysis might be interpreted as a 'quasi-placebo' test. Indeed, at highway stations, no price effect is found.

Fifthly, based on stations' names and brand affiliations, we define a group of stations that likely do not sell any alcohol-related products at night, so that the policy should not affect the outcome. In terms of the model in Section 1.3, this reflects a situation where no consumer is interested in alcohol or where alcohol does not give any utility to consumers. For means of econometric power, the respective group of stations pools supermarket stations, unmanned stations and car dealer stations. Supermarket stations most of the time do not have a shop at all as they typically are owned by the supermarket nearby (Haucap et al. (2017*b*)). Though, supermarkets are closed during the nightly prohibition (10pm to 5am), so that supermarket stations are not affected by the policy lifting. Unmanned gasoline stations (e.g., by the brand AVIA Xpress) do not operate a shop (at night). Similarly, car dealer stations' main purpose is to provide fuel for the main business. As expected, we find that such stations, indeed, do not change prices in response to the policy (s. fifth panel of Figure 1.3).

Sixthly, we investigate whether the price effect is mitigated by the fact whether stations are located in urban counties or the periphery. We follow the county-level definition of urbanity by *Federal Institute for Research on Building, Urban Affairs and Spatial Development.* We find stations in urban vicinities to increase prices more strongly (s. sixth panel of Figure 1.3). This likely reflects the higher share of youths in urban regions, which Marcus and Siedler (2015) found to increase their alcohol demand. Hence, in urban stations, more consumers should receive utility from the newly available product after the prohibition lifting.

In the last panel of Figure 1.3, we study heterogeneity in the local share of youths (18-25 year-olds). This traces back to Marcus and Siedler (2015), who find that the discussed alcohol prohibition especially reduced alcohol binge consumption among young adults. We investigate whether a higher share of youths proxies a demand shock for gasoline as well. With regard to the model, youths might reflect a consumer group, who is aware of the alcohol product (high λ in the model above). As they gain most from the availability of alcohol, the alcohol-driven demand shift should cause higher prices for stations with a high local share of youths (s. Result 2 in Section 1.3). Our estimates do not reveal a clear treatment effect heterogeneity when comparing stations from municipalities above and below the median youth share. Though, when zooming in on the heterogeneity of the youth share more intensively, a clear relation between a higher youth share and a higher treatment effect is evident. For example, see Figure 1.A.1 in the appendix for treatment effect heterogeneity across terciles and quartiles of the distribution.

Note that we, as a robustness check, also ran our heterogeneity analysis in a single regression instead of separate regressions. This should ensure that the different heterogeneity results are not driven by one and the same factor which correlates with several station characteristics. Table 1.A.2 in the appendix presents these results. Qualitatively our results do not change. Especially stations with few competitors at night and in municipalities with a high share of youths experience higher treatment effects. Also, small assortment stations increase prices more strongly.

Station Activity. As we find that gasoline prices at stations in Baden-Wuerttemberg during the prohibition have been lower, there likely is an unambiguous effect on the overall revenues of stations: Alcohol revenues vanish and gasoline prices drop. Hence, it is a natural question whether some stations change how actively they participate in the market in response to the policy lifting.

To study stations' activity, we use the real-time price data to elicit whether a gasoline station in a certain week changed prices at night or not. If stations change prices, this will be indicative of whether they open at night. Due to data availability, we cannot fully exclude that effects on price changes are shaped by Edgeworth cycle adaptions due to the policy instead of operating times. Though, in the appendix, we provide some evidence in Table 1.A.3 on whether gasoline stations in Baden-Wuerttemberg show different Edgeworth cycle characteristics after the policy lifting. The number of price changes over the day as well as cycling frequency and asymmetry remain unaffected.

We determine whether a station has changed its price between 10pm and 5am and, for a second measure, whether there have been changes between 11pm and 4am. We apply a standard dynamic DID estimator in a two-way fixed effects model to study stations' propensity to operate at night. Again, flat pre-trends will be indicative of whether the parallel trend assumption holds:

$$1[Active\ at\ Night]_{sw} = \alpha_s + \lambda_w + \sum_{t=\underline{\tau}, t\neq -1,-2}^{\bar{\tau}} \gamma_t 1[BW_s \times Lifting_{w-t}] + \epsilon_{sw} \tag{1.4}$$

$1[Active\ at\ Night]_{sw}$ is a dummy which will turn one if a station s has operated at night in week w. We apply two definitions for this outcome: Firstly, the variable will turn one if a station is active/changes the price at least once a week between 10pm to 5am. Secondly, the variable will turn one if a station is active/changes the price between 11pm and 4am at least once a week.

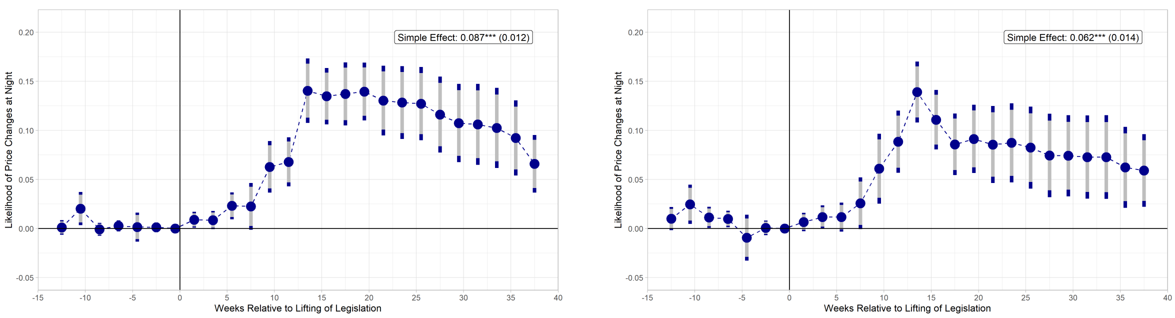


Figure 1.4: Dynamic Effects on Likelihood to be Active at Night

Note: This plot gives dynamic estimates of the leads and lags from equation (1.4). The left plot defines $1[Active\ at\ Night]_{sw}$ with changing prices between 10pm and 5am, the right plot takes a more restricting definition of price changes between 11pm and 4am. Standard errors are clustered at the county level. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients from a linear probability model.

Figure 1.4 gives the dynamic estimates for both outcomes. It appears that the share of stations being active at night increases substantially after the lifting of the prohibition. In fact, stations in Baden-Wuerttemberg are 8.7 percentage points (or 10% respectively) more likely to operate/change prices at some point between 10pm and 5am than when the prohibition was active. In contrast to the price effect, which arises after 5-7 weeks, the reaction in night activity takes about twice as long until reaching a constant treatment effect level. This is very much in line with lower menu costs for price level changes than structural changes in a station's activity at night.

When investigating heterogeneous responses across stations with small or large assortment, we find heterogeneity, which corresponds to the price effects found above. Stations with a small assortment typically sell fewer products, so that a restriction on alcohol might hit them more strongly. Indeed, we find that such stations react more pronouncedly in activity during prohibition hours (s. Figure 1.5). We also checked again, whether highway stations do not react to the policy in means of nightly activity and, indeed, that is observed.

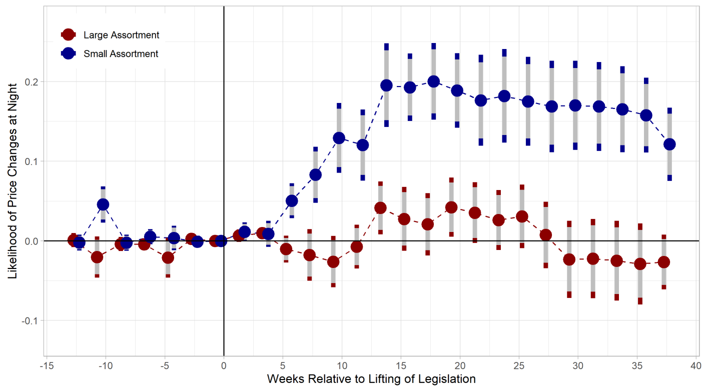


Figure 1.5: Dynamic Effects on Station Activity: Heterogeneity Along Assortment Variety

Note: This plot gives dynamic estimates of the leads and lags from equation (1.4) for two subsamples of stations with heterogeneous store assortment. The outcomes $1[Active\ at\ Night]_{sw}$ is defined as the weekly share on which prices have been changed between 10pm and 5am. Standard errors are clustered at the county level. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients from a linear probability model.

A concern is that price changes might not perfectly reflect opening hours. For example, stations that open at night but only start to change prices after the prohibition lifting, are implicitly understood to extend opening hours due to the policy. Hence, this would likely upward bias the estimated treatment effect. Therefore, we provide additional robustness checks on the effect of opening hours (s. Section 1.7 in the appendix for an in-depth analysis). Using historical opening hours for a subset of gasoline stations ($\approx$ 25% of all stations), which we obtained from the internet archive `web.archive.org`, we find opening hour reactions in line with our results above. Again opening hours are increased significantly in Baden-Wuerttemberg after the policy lifting - especially at stations with smaller shop assortment. Though, our robustness check identifies smaller treatment effects. This is likely due to the potential upward bias in the analysis based on price changes as explained above.

Finally, the extended opening hours likely cause our baseline price effect to be downward-biased as more competitors have been found to correlate with lower prices in gasoline markets (Haucap et al. (2017*a*), Martin (2023), Pennerstorfer et al. (2020)). In the appendix, we show how a change in the number of nighttime competitors affects nighttime prices. Table 1.A.4 reports the results of an interaction term analysis in columns (1) and (2). We find that the treatment effect is larger in concentrated markets. In Figure 1.A.3, we exploit the staggered timing of competitors' nighttime entry across incumbents and show that nighttime entry in a 1km radius decreases prices by up to 1 Eurocent/l[12]. The effect size is very similar to Fischer et al. (2023) who estimate the causal effect of station entry on incumbent prices to be around 0.5ct/l. This indicates, that, indeed, our baseline results are downward-biased.

As nighttime entry decreases prices by up to 1 Eurocent/l, it absorbs the policy-induced price effect completely in markets where nighttime entry takes place. However, nighttime entry is costly and might not be possible or profitable in all markets, so that positive price effects remain in the majority of markets, which are not entered. This leads to the on average positive price effect of the policy found above. In columns (3) and (4) of Table 1.A.4, we try to quantify by how much nighttime entry decreases the price effect which would have been observed absent nighttime entry. For this, we include the entry of competitors as 'bad control' in the price regressions. We show that the price effect changes only slightly in comparison to the estimated baseline effect. This indicates that entry only marginally decreases the policy's average price effect.

Traffic Flow Analysis. To better understand the mechanism underlying the observed price effects, we study traffic flow reactions to the policy. The analysis is twofold: First, we analyse whether nightly traffic increases in response to the policy lifting in Baden-Wuerttemberg and especially near open gasoline stations. Secondly, we study how traffic at the federal state's border is affected by the shock.

We start by running the triple difference-in-differences regression from above on the logged number of counted cars for traffic counters near gasoline stations open at night ($\leq$ 2km linear distance). Figure 1.6's blue estimates report the dynamic effect of the policy lifting on traffic near gasoline stations in Baden-Wuerttemberg in four-week bins. After the policy lifting, nightly traffic in Baden-Wuerttemberg persistently increases by up to 5-10%. This is indicative of more cars traveling near and, hence, likely also to gasoline stations. This is in line with a demand expansion through the *service quality* channel as alcohol is available after the policy. To show that this effect really reflects an increasing interest in gasoline stations, we run this analysis separately for groups of traffic counters that have different distances to the nearest open gasoline station. The traffic effect should be highest for counters near gasoline stations if traffic increases really relate to more visits to gasoline stations. Figure 1.7, indeed, shows that this is the case. While traffic in Baden-Wuerttemberg overall increases by around 5%, this effect is strongest for counters right next to gasoline stations ($\leq$ 1km linear distance). There is no significant effect on traffic flows for counters more than 2km away from open gasoline stations. We take this as support for our demand expansion channel.

In addition, Figure 1.A.2 in the appendix reveals that the increase in traffic is especially high for traffic counters in municipalities with a high share of youths. This fits the story that especially youths respond to the policy change.

We further study border traffic. Before the policy lifting, consumers living in Baden-

[12]Similar procedures can be found in the reduced-form entry literature as in Arcidiacono et al. (2020), Goolsbee and Syverson (2008) and Matsa (2011).

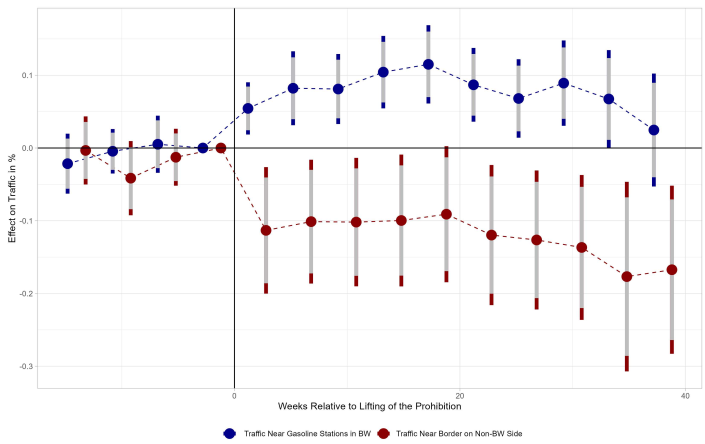


Figure 1.6: Dynamic Effects on Traffic Flows

Note: This plot gives dynamic estimates of the leads and lags from equation (1.4) where the outcome variable is logged traffic flows. The blue estimates give the effect of the policy on traffic counts in Baden-Wuerttemberg near gasoline stations ($\leq$ 2km linear distance) in a subsample of traffic counters of maximum 2km linear distance to gasoline stations open at night. The red estimates give the effect of the policy of traffic counts near the border ($\leq$ 2km linear distance) to Baden-Wuerttemberg at non-Baden-Wuerttemberg counters in a subsample of non-Baden-Wuerttemberg traffic counters. To account for the logarithm of very few zero traffic observations, we use the hyperbolic sine transformation of the outcome variable. Standard errors are clustered at the county level. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients.

Wuerttemberg had to leave the federal state to get alcohol at night at off-premise locations. This border traffic should have been reduced after the policy lifting. For this, we compare traffic at traffic counters outside of Baden-Wuerttemberg but near the border ($\leq$ 2km linear distance) to all other non-Baden-Wuerttemberg traffic counters before and after the policy. Figure 1.6's red estimates report the results of the triple difference-in-differences regression. Indeed, traffic near the border to Baden-Wuerttemberg but outside of Baden-Wuerttemberg falls in response to the policy. This can be interpreted as a demand shift to gasoline stations in Baden-Wuerttemberg. Also, this result indicates that alcohol consumption has a sufficiently high value to consumers to induce border travel. Note that we also tried out other distance thresholds up to 5km distance to the border and our results qualitatively remain the same.

We, further, reproduce the heterogeneity analysis from Figure 1.3 with traffic flows as an outcome to support the mechanisms described above. To conduct heterogeneity analyses along gasoline station characteristics (brand, shop size, etc.), we match counters to the nearest station. The results in Figure 1.A.4 in the appendix show that traffic increases more strongly at counters with many stations nearby and also is stronger in urban areas with a high youth share.

We complement the traffic data results on a demand expansion mechanism with an analysis of geo-coded traffic accidents with personal damage in Germany[13]. In Table 1.A.5, we, at the

[13]The data comes from the 'Unfallatlas' (`https://unfallatlas.statistikportal.de/`) of the Federal Statistical Office and the Statistical Offices of the German States and covers traffic accidents with personal damage for 12 out of 16 federal states.

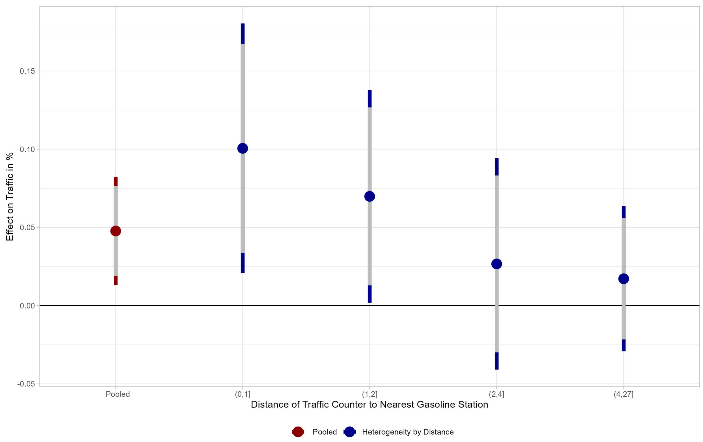


Figure 1.7: Effects on Traffic Flow by Counter Distance to Station

Note: This plot gives the estimates from the static version of the equation (1.4) where the outcome variable is logged traffic flows. Stations are grouped by the minimum distance to a gasoline station open at night. To account for the logarithm of very few zero traffic observations, we use the hyperbolic sine transformation of the outcome variable. Standard errors are clustered at the county level. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients.

extensive margin, do not find an effect of the policy lifting on the overall number of accidents with personal damage in Baden-Wuerttemberg.[14] However, we show that the likelihood of accidents being very near ($\leq$ 1km) to open gasoline stations increases by 3% after the policy lifting. On average, the distance of accidents to the nearest gasoline station at night decreases by 7% after the policy lifting. This shows that traffic flows likely shift towards areas surrounding gasoline stations.

Bite of the Policy. To quantify the consequences of the policy for gas stations as well as consumers, it is not sufficient to show that the price effect is around 5% of an average station's margin. We need to understand how many consumers visit gasoline stations at night. To approximate *daytime-specific* demand, we rely on *Google Popularity* data[15], which we scraped for all stations available once in July 2019 ($\approx$ 85% of all German stations). Figure 1.8 plots the average distribution of gas station visits over the course of the day. Non-negligible 7-8% of visits lie in the treatment time between 10pm and 5am[16].

Moreover, stations do not only use revenues from gasoline sales during the prohibition but also lose alcohol revenues. Industry surveys (Scope Ratings (2018)) show that the annual alcohol revenues of an average station are approximately 100,000 Euro.

Furthermore, consumers potentially switching away from stations, which increase prices more strongly after policy lifting, are a concern when discussing the exposure of consumers to the

[14]This is in line with the results in Baueml et al. (2023) who do not find the policy's introduction in 2010 to affect alcohol-related traffic accidents.

[15]On *Google Maps*, it is reported how crowded and popular a business is for every hour of the day. Popularity is based on measures such as mobile phone mobility and traffic and is reported in an index between 0 and 100 at the station level.

[16]In a telephone survey of the German Ministry for Economic Affairs and Energy from 2016, the share of respondents who fuel at this time is of similar magnitude (Bundesregierung (2018)).

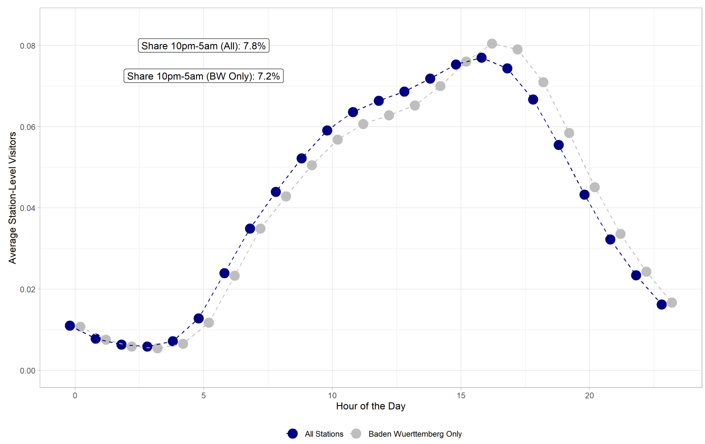


Figure 1.8: Average Share of *Google Popularity* Across Stations

Note: This plot gives the distribution of *Google Popularity* as a proxy of demand over the course of the day. This is separately reported for all stations in Germany and Baden-Wuerttemberg only. As in the rest of the paper, we only consider those stations which open 24/7, i.e. those for which popularity data is available for all hours of the week.

policy. Especially informed consumers would not be affected by the policy then and distributional implications would arise. While we do not observe actual transactions - so where consumers fuel - we can show that consumers can hardly avoid being affected by the policy effect as long as the policy shifts the complete price distribution in a first-order stochastic dominance manner. Then, consumers at all percentiles of the distribution are affected.

In appendix 1.7, using the method of Chernozhukov et al. (2013), we show that the prohibition lifting indeed shifts the price distribution in a first-order stochastic manner. As prices at all quantiles increase, consumers can hardly avoid the exposure to the policy's price effect.

Other Mechanisms. While we argue for a trade-off between a *service quality*-induced demand expansion channel and a *cross-subsidization* channel, other policy-induced changes in consumer or firm behavior could potentially explain the observed price effects. Examples are a policy-induced change in consumer information about prices or the demand elasticity. We discuss relevant, alternative mechanisms subsequently to justify that they do not drive the findings.

It is a concern that the composition of consumers changes in response to the treatment. Firstly, some consumers might only visit stations to buy alcohol and do not anticipate buying gasoline. These consumers do not compare gasoline prices in advance and, hence, the declining relevance of competition nearby might rationalize the observed price increase. Though, this then should only bite in less concentrated markets. To the contrary, our heterogeneity analysis in Figure 1.3 reveals the opposite as the price effect is especially driven by stations with few competitors.

Secondly, the share of informed consumers might change with the treatment. For example, fewer people might compare gasoline prices as some consumers mainly come for alcohol and do not anticipate buying gasoline. This could result in increasing nighttime prices, too. In this case, the share of consumers, which do not consider gasoline prices at competing stations, increases. Such consumers behave like non-shoppers in Varian (1980). While we have no explicit information about changes in the information structure of consumers, we can test in the data whether the effect of a change in consumer information on market-level prices reflects the theoretical predictions in Varian (1980). We do this in Section 1.7 of the appendix in very detail and show that observed changes in the estimated upper bound of the market-level, monthly price distribution in response to the treatment suggest a change in the observed valuation of a purchase instead of a change in the share of informed consumers. We interpret this in favor of the outlined *service quality* channel.

We, moreover, test whether consumer frictions at night develop differently in Baden-Wuerttemberg after the policy lifting. For this, we extend the rank reversal test in Chandra and Tappata (2011) to a difference-in-differences setup. The test's intuition is that rank reversals of a station couple's prices over time in a homogenous product market arise from consumer frictions. We do not find that rank reversals become more likely in Baden-Wuerttemberg after the policy lifting. This contradicts the argument of changing consumer information in response to the policy. See Section 1.7 for more precise explanations and results of this analysis.

Thirdly, nighttime consumers might be less price-sensitive and less price-elastic with respect to the gasoline price after the policy lifting - beyond the discussed changes in information and competition relevance above. A less elastic cross-price elasticity would imply that the relevance of competitor prices decreases. In the most extreme case, stations become quasi-monopolies and prices of neighboring stations are not strategic responses to each other anymore. Hence, a less elastic cross-price elasticity likely results in a weaker price comovement. We empirically investigate whether price comovement of neighboring stations changes with the policy lifting by comparing the correlation of prices between couples in Baden-Wuerttemberg and other states over time in a difference-in-differences setup. Section 1.7 of the appendix gives a detailed explanation of the empirical analysis and the results. We do not find any evidence for a change in the strength of price comovement (s. Table 1.F.1). Hence, we interpret this as evidence for no change in the demand elasticity as the driving mechanism behind the observed price effects.

Fourthly, the treatment might have affected Edgeworth cycle characteristics in Baden-Wuerttemberg which could explain the observed price effects. For example, the policy-induced market entry at night might change the cycle structure as other papers have found the competition to shape cycles (Noel (2007), Siekmann (2017)). Table 1.A.3 shows that there are no significant changes in typical cycle characteristics related to the policy lifting.

Fifthly, one might argue that alcohol could also be the bait for gasoline, which would explain why introduction of alcohol sales increases gasoline prices. Though, this is unlikely for three reasons: First, the share of consumers not considering to fuel while traveling to a gasoline station for alcohol likely is low. Hence, most consumers willing to fuel will account for gasoline prices. This effectively limits the potential for gasoline price increases. Second, high price transparency for gasoline through price apps, websites, and price signs in front of gasoline stations limits the extent to which firms can change the add-on price. Pure gasoline consumers would then switch away. Lastly, the prohibition hours at night lie in the time period of the day in which intra-day gasoline price cycles peaked in Germany at this time. Hence, consumers can become better off by avoiding high add-on prices by switching intertemporally. This is not the case for alcohol prices, which do not vary between daytimes.

Lastly, our baseline, reduced-form results leave it open whether the dominated *cross-subsidization* channel exists at all. However, our heterogeneity analysis for shops with a smaller and larger assortment reveals smaller price effects for larger shops. This likely reflects the higher incentive to cross-subsidize for shops with larger assortments.

1.6 Robustness Checks

The price effect of the legislation lifting may be especially high if consumers are aware of alcohol again being available at gasoline stations. This could reflect a stronger demand shock. Hence, consumer awareness might be essential. Even though consumers might be implicitly steered through shops, some consumers actively decide to visit gasoline stations to buy products in the shop. While Baueml et al. (2023) provide survey evidence that people were aware of the prohibition, there is no evidence on the familiarity with the policy lifting. We investigate consumer awareness by studying search queries in *Google Trends*, which documents standardized search frequencies for keywords in the search engine. Google searches have been used in previous literature to study policy awareness or agents' behavior as well (Garthwaite et al. (2014), Isphording et al. (2021), Lichter and Schiprowski (2021)). Google documents weekly search frequencies for given phrases at the state level. We gather time series of search frequencies for 23 policy-related keywords through the API of the R package *gtrends* at the keyword-state level and estimate a dynamic DID setup with the standardized search frequency across states and over time. The respective regression is as follows:

$$Search_{kfw} = \theta_f + \eta_{kw} + \sum_{t=\underline{\tau},t\neq-1}^{\bar{\tau}} \phi_t 1[BW_f \times Lifting_{w-t}] + \epsilon_{kfw} \tag{1.5}$$

where $Search_{kfw}$ is the standardized search frequency for keyword k in federal state s and week w. θ_f and η_{kw} are state and keyword-week fixed effects, so that identification stems from within-keyword changes in the search frequency over time. $\sum_{t=\underline{\tau},t\neq-1}^{\bar{\tau}} \phi_t 1[BW_f \times Lifting_{w-t}]$ are dummies which will be one if f =Baden-Wuerttemberg and if the prohibition's lifting is t periods ago. Hence, ϕ_t are the coefficients of interest and document whether there have been more or fewer search frequencies in comparison to the control states relative to one period before the treatment.

We include searches related to the policy, for example, 'Alcohol Selling Prohibition Baden-Wuerttemberg' (in German: Alkoholverkaufsverbot Baden-Württemberg),'Gasoline Station' (Tankstelle), 'Alcohol Gasoline Station' (Alkohol Tankstelle), 'Gasoline Station Opening Hours' (Tankstelle Öffnungszeiten), 'Baden-Wuerttemberg Alcohol' (Baden-Württemberg Alkohol) and others.

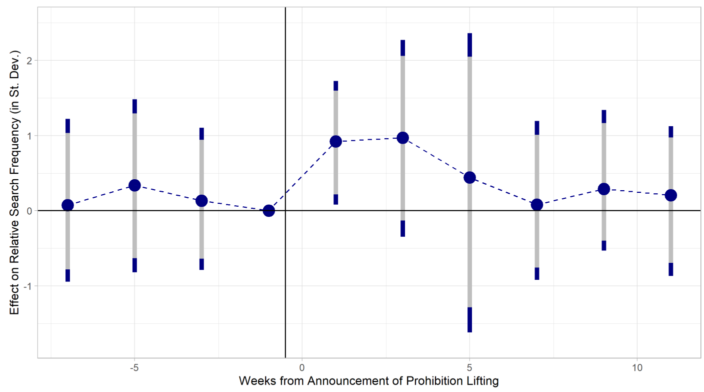


Figure 1.9: Dynamic Effects on Policy Awareness

Note: This plot gives dynamic estimates of the leads and lags for the DID model in equation (1.5). The outcome is the standardized search frequency. Standard errors are clustered at the state level ($n = 16$). We apply the wild-bootstrap inference with 499 repetitions to account for the small number of clusters. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients. $N = 7,360$ observations across 20 weeks, 16 federal states, and 23 keywords.

Figure 1.9 gives the effect of the lifting's announcement - around one week before the actual lifting - on the search frequencies for related keywords in Baden-Wuerttemberg relative to the other German states. As can be seen, in Baden-Wuerttemberg, the policy receives attention right after the policy announcement and the policy lifting. No anticipatory awareness is evident. The additional search frequency of up to a standard deviation only holds for a few weeks when the search intensity drops to the former level again. This is in line with attention at the time of the shock. This supports consumer awareness of the policy change.

Beyond studying policy awareness, we implemented further robustness checks. First, we combined our difference-in-differences regression with a propensity score matching to eliminate differences across treatment and control stations in observable characteristics. We matched treated stations to control stations within the last pre-treatment period. We match treated stations to their nearest neighbor without replacement. We drop treated stations that did not set a price in the pre-treatment week ($\approx 10\%$) and arrive at 733 station pairs. Tables 1.A.6 and 1.A.7 show that the sample is balanced after matching and that our regression results do not change qualitatively.

We, further, tested our empirical setup's robustness to including other fixed effects combinations or additional state-level time trends and price effects near state borders in Tables 1.A.8 and 1.A.9. Our results do not change in the presence of the trends and the different fixed effects allocation. Our border analysis further reveals no price effect at the state border. This fits the ex-ante hypothesis that competition and strategic complementarity between treated and untreated stations lead to a null effect when both types of stations are in the near vicinity. We also checked heterogeneity in the treatment effect depending on which region (federal state) of Germany is chosen as a comparison group. Figure 1.A.5 gives the respective estimates and shows that the effect is positive and statistically significant for most of the states as control group. Only, Bavaria and Lower Saxony reveal negative estimates.

Hence, the effect is barely sensitive to specific regions of Germany. Moreover, we provide additional evidence on other inference methods for our baseline estimate in Table 1.A.10. In fact, clustering at the county level is a conservative approach as markets are often defined on the granular municipality level (Pennerstorfer et al. (2020)) or studies cluster at the market level (Assad et al. (2021)).
Lastly, we study how the treatment effect varies when changing the pre- and post-treatment effect window. Table 1.A.11 shows that the treatment effect is quite robust across different window choices.

1.7 Conclusion

This paper examines (unintended) spillover effects of a nightly off-premise prohibition for alcohol sales in Baden-Wuerttemberg, Germany. Applying difference-in-differences setups, we find that gasoline prices in Baden-Wuerttemberg increased by around 0.6 Eurocent/l after the lifting of the prohibition ($\approx 5\%$ of the net margin). We argue that gasoline stations exploit being 'stores of last resort' for alcohol at night. As opportunity costs of fuelling at a different station from where alcohol is purchased are high, alcohol consumers create a demand shock for stations. The effect size increases in the absence of many competitors and is especially high at stations with small shop assortments.
Implications for policymakers arise. Our analysis shows that gasoline stations rely on multiple revenue channels and strategically consider their price interactions. Product variety as means of add-on quality is positively priced in gasoline prices. Stations do not cross-subsidize between a transparently priced product (gasoline) and a less transparently priced product (alcohol). These findings have implications for market definition, which - up to now - mostly is limited to gasoline businesses themselves in the literature. Price relations between gasoline and consumables though indicate that competition on shop products (for example with supermarkets) may show price effects at the pump as well. Further evidence on market delineation and spillovers from shop-related regulation on gasoline prices could give new insights to those questions.
Second, our results hint at distributional effects which will arise if consumers are heterogeneously informed. It may even be that commonly applied price transparency regulations, which make gasoline prices more salient, leverage the mismatch of uninformed consumers and high add-on quality stations.

Appendix A: Figures and Tables

Table 1.A.1: Descriptive Statistics: Traffic

Statistic	Units	Control (Pre-Lifting) (1)	BW (2)	Δ (p-value) (3)
Outcomes				
ln(Traffic (Day))	#	11.902	11.872	0.81
ln(Traffic (Night))	#	9.527	9.584	0.71
Location				
Distance to Station Open at Night	km	3.802	3.175	0.06*
ln(Distance to State Border)	#	4.972	3.712	0.00***

Note: This table compares descriptive statistics of counters in Baden-Wuerttemberg with counters outside of Baden-Wuerttemberg (both pre-treatment). The p-values come from linear regressions of the respective outcome on an intercept and a dummy for Baden-Wuerttemberg where we implement standard errors clustered at the county level.

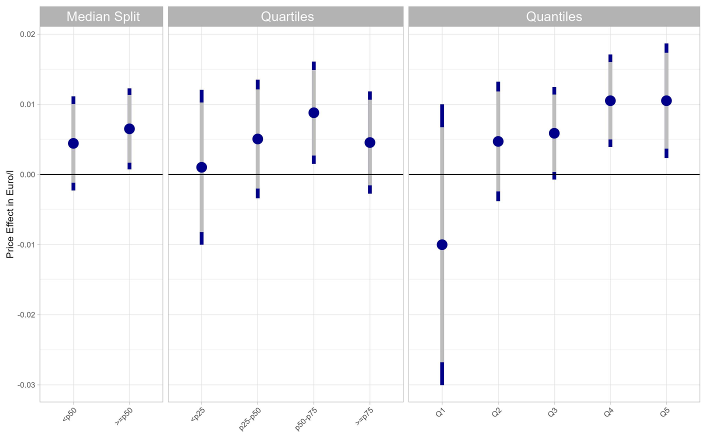


Figure 1.A.1: Price Effect: Heterogeneity Along 'Youth Share' Distribution

Note: Heterogeneity analysis based on sample splits along the distribution of the variable 'Youth Share'. 90% and 95% confidence bands are reported. Standard errors are clustered at the county level.

Table 1.A.2: Heterogeneity Analysis: Robustness Check

	Gasoline Price in Euro/l	
	(1)	(2)
BW×Post×1[Street Station]	−0.0064	
	(0.0079)	
BW×Post×1[Urban]	0.0054	
	(0.0051)	
BW×Post×1[Below Median Competition]	0.0064***	
	(0.0024)	
BW×Post×1[Below Median Youth Share]	−0.0048*	
	(0.0027)	
BW×Post×1[Large Assortment]	−0.0127***	
	(0.0046)	
BW×Post×1[Oligopolistic]	−0.0024	
	(0.0042)	
BW×Post×1[No Shop Sales]	−0.0102*	
	(0.0062)	
BW×Post×Night×1[Street Station]		−0.0059
		(0.0076)
BW×Post×Night×1[Urban]		0.0056
		(0.0050)
BW×Post×Night×1[Below Median Competition]		0.0063**
		(0.0024)
BW×Post×Night×1[Below Median Youth Share]		−0.0045*
		(0.0026)
BW×Post×Night×1[Large Assortment]		−0.0131***
		(0.0046)
BW×Post×Night×1[Oligopolistic]		−0.0024
		(0.0042)
BW×Post×Night×1[No Shop Sales]		−0.0110*
		(0.0062)
Sample	Night Prices	All Prices
Observations	296,598	593,193
Adjusted R^2	0.894	0.911

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup extends the regression equation from the 'Data and Empirical Strategy' section by additional interactions. Model (1) only uses night prices and a triple difference-in-differences estimator, while model (2) uses quadruple interactions to extend the baseline triple difference-in-differences estimator to account for effect heterogeneity. Other interactions not reported in the regression table.

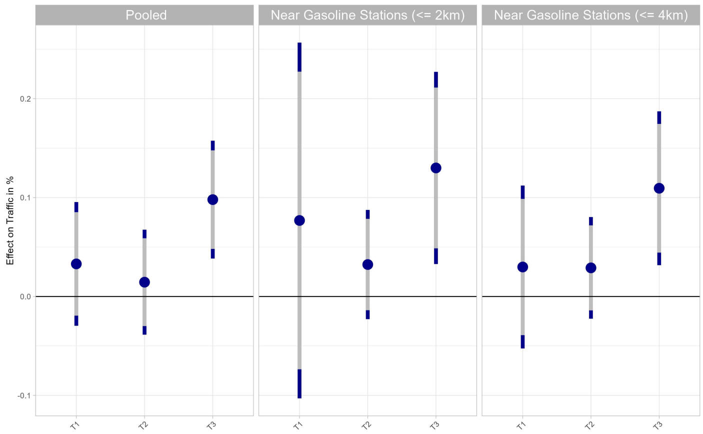


Figure 1.A.2: Heterogeneity in Traffic Response Along Youth Share Distribution

Note: This plot gives the estimates from the triple DiD model presented in the Section 'Data and Empirical Strategy' with trafficflows as outcome. The analysis is run for all stations, only stations in a radius of 2km linear distance to gasoline stations open at night or a 4km radius. 90 and 95% confidence bands are reported. Standard errors are clustered at the county level.

Table 1.A.3: Edgeworth Cycle Characteristics

	Median Price Change	ln(# Price Changes)	Price Spread
	(1)	(2)	(3)
BW×Post	0.0004	0.0225	0.0018
	(0.0003)	(0.0152)	(00017)
Approach	DID	DID	DID
Observations	2,155,817	2,156,356	2,118,970
Adjusted R^2	0.189	0.753	0.591

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows a simple DID.

Table 1.A.4: Regressions on Mitigating Entry Effect

	Gasoline Price in Euro/l			
	(1)	(2)	(3)	(4)
BW×Post	0.0097***		0.0082***	
	(0.0027)		(0.0024)	
BW×Post×(# Competitors ∈ [0,1] km)	−0.0043***	−0.0039**		
	(0.0015)	(0.0015)		
BW×Post×(# Competitors ∈ (1,2] km)	−0.0004	−0.0002		
	(0.0011)	(0.0011)		
# New Competitors Active ∈ [0,1] km			−0.0101***	−0.0091***
			(0.0033)	(0.0031)
# New Competitors Active ∈ (1,2] km			−0.0042	−0.0033
			(0.0028)	(0.0028)
Station FE	✓	✓	✓	✓
Week FE	✓	×	✓	×
State×Week FE	×	✓	×	✓
Observations	296,598	296,598	296,598	296,598
Adjusted R^2	0.878	0.882	0.876	0.880

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression equation is a triple difference-in-differences regression for nighttime prices comparing prices across federal states, before and after the policy and across competition environments. The other interaction terms of triple difference-in-differences estimator in columns (1) and (2) are omitted.

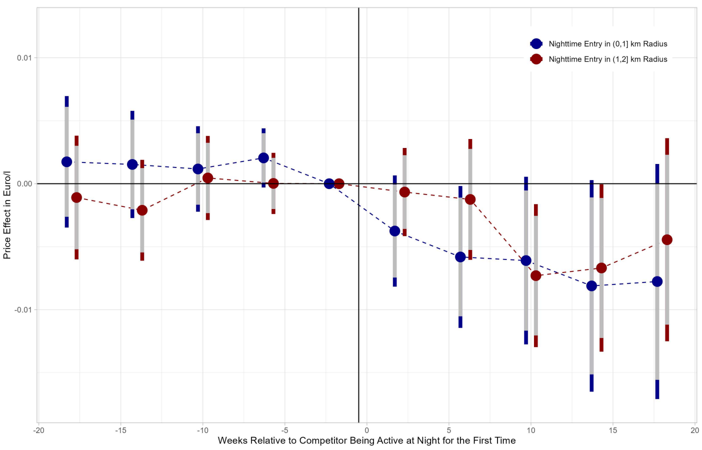


Figure 1.A.3: Effects of Nighttime Entry on Prices

Note: This plot gives estimates from an event study regression of nightly prices on leads and lags of the nighttime entry of competitors in a 1km or 1km-2km radius around incumbents. Station and state-week fixed effects are included. Nighttime entry of stations is identified in the week after which a station operates two consecutives weeks at night for the first time. Endpoints are binned and not reported due to an unbalanced panel in event time (Fuest et al. (2018)). Standard errors are clustered at the county level. The exact timing of entry is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients.

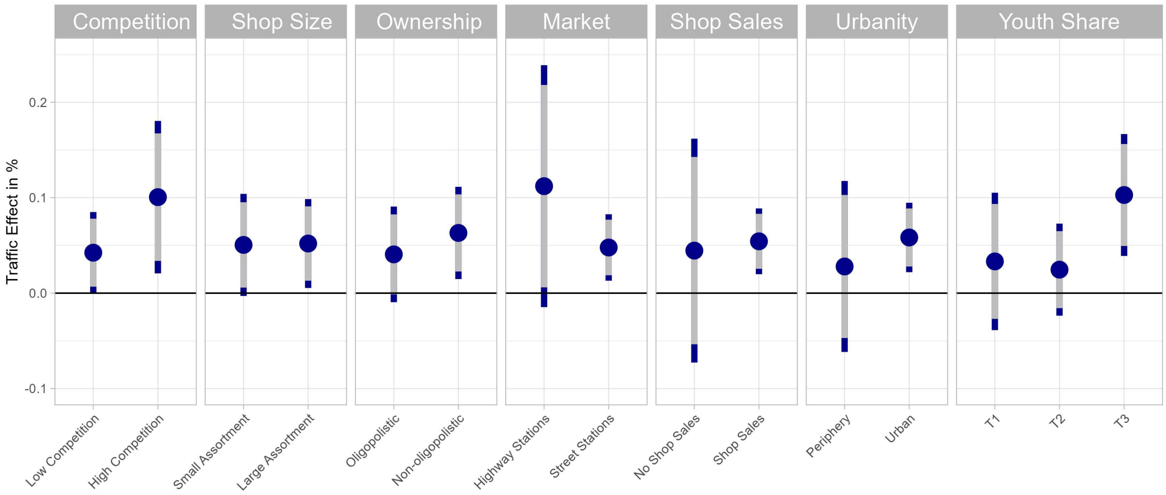


Figure 1.A.4: Heterogeneity Analyses: Traffic

Note: This plot gives the treatment effect of the policy on traffic flows for subsamples along counter and station characteristics. The y-axis documents the effect size in %, the x-axis gives the respective subsample. 90% and 95% confidence bands are reported. Standard errors are clustered at the county level. To be able to conduct heterogeneity analyses along station characteristics, we match counters to the nearest stations operating 24/7. We only include counters in the analyses that are closer than 5km to a gasoline station.

Table 1.A.5: Policy Lifting's Effect on Accidents

	ln(# Accidents)			1[Distance Station $\leq$ 1]			log(Distance Station)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BW×Post×Night	−0.017			0.031**			−0.069**		
	(0.022)			(0.015)			(0.035)		
BW×Post		−0.027	−0.043		0.004	0.036**		−0.003	−0.069**
		(0.023)	(0.030)		(0.004)	(0.014)		(0.010)	(0.034)
County FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Month×Hour FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Approach	TDID	Only Day	Only Night	TDID	Only Day	Only Night	TDID	Only Day	Only Night
Observations	182,016	128,928	53,088	393,445	368,544	24,901	393,445	368,544	24,901
Adjusted R^2	0.512	0.426	0.241	0.170	0.158	0.187	0.169	0.148	0.273

Note: *p<0.1; **p<0.05; ***p<0.01. Regressions (4) to (9) are at the individual accident-level and hence also include controls for whether the accident included bicycles, motorbikes and pedestrians. Regressions (1) to (3) are at the county-month-hour level. The distance to the nearest *open* station depends on the time of the day. The post-treatment period is the first full month after the policy lifting and beyond. Standard errors are clustered at the county level.

Table 1.A.6: Robustness Checks: Propensity Score Matching - Balancing Condition

	Before Matching			After Matching		
	Control	BW	Δ (p-value)	Control	BW	Δ (p-value)
Outcomes						
ln[E5 Gasoline Price (Day)]	0.320	0.314	0.00***	0.315	0.314	0.74
ln[E5 Gasoline Price (Night)]	0.368	0.364	0.03**	0.364	0.364	0.99
ln[Margin (Day)]	-2.385	-2.473	0.00***	-2.463	-2.473	0.59
ln[Margin (Night)]	-1.943	-1.989	0.01**	-1.988	-1.989	0.97
Competition						
# Competitors 0.5km Radius (Day)	0.471	0.452	0.51	0.467	0.452	0.70
# Competitors 0.5km Radius (Night)	0.258	0.229	0.17	0.231	0.229	0.96
# Competitors 1km Radius (Day)	1.078	1.097	0.71	1.079	1.097	0.79
# Competitors 1km Radius (Night)	0.546	0.538	0.79	0.502	0.538	0.41
Stations Characteristics						
Share of Youths (18-25 y.o., Munic. Level)	0.075	0.083	0.00***	0.083	0.083	0.99
Premium Station	0.439	0.415	0.21	0.394	0.415	0.43
Oligopolistic Station	0.373	0.276	0.00***	0.247	0.276	0.21

Note: *p<0.1; **p<0.05; ***p<0.01. Matching was done in a sample of observations from the last pre-treatment week only. Matching was conducted with nearest neighbor matching without replacement. Only stations, which set a price (i.e., which were active) in the respective week, were included in the matching regression. Only observations with positive margins included in the matching regression.

Table 1.A.7: Robustness Checks: Propensity Score Matching - DiD Results

	Gasoline Price in Euro/l			ln(Gross Margin)
	(1)	(2)	(3)	(4)
BW×Night×Post	0.0051**			0.0702***
	(0.0025)			(0.0219)
BW×Post		0.0068**	0.0018	
		(0.0028)	(0.0020)	
Approach	TDID	DID	DID	TDID
Sample	Baseline	Only Night	Only Day	Baseline
Observations	147,818	73,909	73,909	147,796
Adjusted R^2	0.887	0.865	0.952	0.768

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows the regression equation from the 'Data and Empirical Strategy' section. The sample is based on a propensity score matching estimator with nearest-neighbor matching without replacement within the last pre-treatment period.

Table 1.A.8: Robustness Checks: TDID Setup

	Gasoline Price in Euro/l						
	(1)	(2)	(3)	(Baseline)	(5)	(6)	(7)
BW×Night×Post	0.0055**	0.0055**	0.0055**	0.0056**	0.0056**	0.0056**	0.0056**
	(0.0022)	(0.0022)	(0.0022)	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Approach	TDID	TDID	TDID	TDID	TDID	TDID	TDID
BW dummy	✓	×	×	×	×	×	×
Post dummy	✓	✓	×	×	×	×	×
Night dummy	✓	✓	✓	×	×	×	×
BW×Post	✓	✓	✓	✓	×	×	×
BW×Night	✓	✓	✓	✓	✓	×	×
Post×Night	✓	✓	✓	×	×	×	×
Station FE	×	✓	✓	✓	✓	✓	✓
Week FE	×	×	✓	✓	✓	✓	✓
Night×Week FE	×	×	×	✓	✓	✓	✓
BW×Week FE	×	×	×	×	✓	✓	✓
Night×Station FE	×	×	×	×	×	✓	✓
State Trends	×	×	×	×	×	×	✓
Observations	593,193	593,193	593,193	593,193	593,193	593,193	593,193
Adjusted R^2	0.072	0.529	0.868	0.889	0.890	0.912	0.914

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows the regression equation from the 'Data and Empirical Strategy' section. The models provide different specifications of a TDID setup.

Table 1.A.9: Robustness Checks: State Border

	Gasoline Price in Euro/l	
	(1)	(2)
BW×Night×Post	−0.0033	−0.0034
	(0.0215)	(0.0077)
Approach	TDID	TDID
Robustness Check	Border (≤ 1km)	Border (≤ 2.5km)
Observations	1,682	7,310
Adjusted R^2	0.874	0.879

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows the regression equation from the 'Data and Empirical Strategy' section. We subsample stations near the policy border.

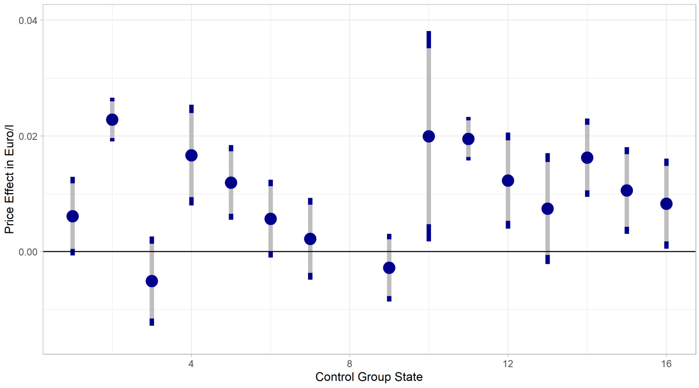


Figure 1.A.5: Treatment Effect for Individual Federal State as Control Group

Note: This plot gives the estimates from the triple DiD model presented in the Section 'Data and Empirical Strategy' for different control groups. In particular, each estimate uses a different federal state as control group. Standard errors are clustered at the county level. We provide 90 and 95% confidence intervals for all coefficients. States are as follows: Schleswig-Holstein (1), Hamburg (2), Lower Saxony (3), Hamburg (4), Northrhine-Westphalia (5), Hesse (6), Rhineland-Palatinate (7), Baden-Wuerttemberg (8), Bavaria (9), Saarland (10), Berlin (11), Brandenburg (12), Mecklenburg-Hither Pomerania (13), Saxony (14), Saxony-Anhalt (15), Thuringia (16).

Table 1.A.10: Inference of Baseline Regression

Coefficient Baseline	0.0056
P-Value	
One-Way Clustering	
Station Level (Baseline)	(0.0014)***
County Level (Baseline)	(0.0022)**
Two-Digit Postcode Level	(0.0029)*
Two-Way Clustering	
Station Level + Week	(0.0005)***
County Level + Week	(0.0009)***
Two-Digit Postcode Level + Week	(0.0010)***
Wild Bootstrap (999 rep.)	
Station Level	(0.0015)***
County Level	(0.0028)**
Two-Digit Postcode Level	(0.0029)*
Cluster Size	
N(Stations)	6,144
N(Counties)	401
N(Postcode Areas)	92
N(Week)	52

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

Table 1.A.11: Different Effect Windows

	Gasoline Price in Euro/l					
	(Baseline)	(2)	(3)	(4)	(5)	(6)
BW×Night×Post	0.0056**	0.0039***	0.0049**	0.0066***	0.0068***	0.0070***
	(0.0023)	(0.0014)	(0.0019)	(0.0020)	(0.0022)	(0.0023)
Effect Window (in Weeks)	[-13, 38]	[-10, 10]	[-10, 20]	[-20, 20]	[-20, 30]	[-20, 40]
Observations	593,193	239,037	353,416	467,986	582,423	696,355
Adjusted R^2	0.889	0.833	0.846	0.833	0.865	0.888

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The outcome variable gives where a station chages the price at least one per week in the time period between 10pm and 5am or 11pm and 4am. The independent variable gives whether a station opens 24/7 in a certain week or not. Observations are at the station×week level.

Appendix B: Quantile Treatment Effects - Robustness Check

To show that the policy lifting shifts the gasoline price distribution in a first-order stochastic manner, i.e. the unconditional quantile treatment effects are positive at all quantiles, we elicit the counterfactual price distribution - so prices in Baden-Wuerttemberg absent the policy lifting after December 08, 2017 - in the style of Chernozhukov et al. (2013). To be precise, we estimate by how much the policy lifting increases/decreases the likelihood of price observations to lie below/above certain price thresholds. Formally, the value of the empirical distribution function (ECDF) of the counterfactual distribution at price p is given through the following 'distribution regression' which estimates the change in the propensity of a price to be below p due to the treatment:

$$1[P_{sw}^{E5} < p] = \beta BW_s \times Post_w + \alpha_s + \lambda_w + e_{sw}. \tag{1.6}$$

The value of the counterfactual ECDF is given by the ECDF of the observed prices minus β. Repeating the procedure for multiple p constructs the full counterfactual distribution.

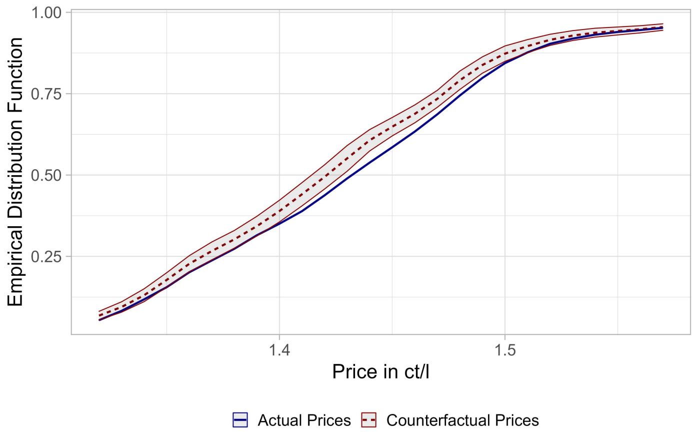


Figure 1.B.1: Distributional Effects of the Policy: Counterfactual Price Distribution

Note: This plot gives the empirical distribution function of observed nighttime post-lifting prices in Baden-Wuerttemberg (blue) and the counterfactual distribution for a scenario without policy lifting (red). The counterfactual distribution comes from distribution regression in the style of Chernozhukov et al. (2013) in one Eurocent/l steps. We provide 95% confidence intervals. Standard errors are clustered at the county level. The distributions are trimmed at the 5th and 95th percentile.

Figure 1.B.1 visually compares observed and counterfactual prices. The policy lifting shifted the price distribution in a first-order stochastic manner to the right. This is indicative of all consumers being affected as the effect is not just driven by one part of the distribution. Instead, consumers at all quantiles of the distribution are affected.

Note that we use weekly average prices and hence we do not fully show that there is no switching opportunity for consumers at a certain point in time which avoids price increases. Admittedly, weekly average prices are indicative of a lack of such switching opportunities. Nevertheless, we encounter this concern by running the same distribution analysis for daily station prices at midnight. We chose midnight as timing as there are barely any price changes after midnight until the end of the prohibition (5am) - s. Figure 1.B.2. Hence, the price distribution at midnight likely reflects the price distribution at 1am, 2am and, hence, most parts of the nightly prohibition period between 10pm and 5am. Results do not change qualitatively (s. Figure 1.B.3).

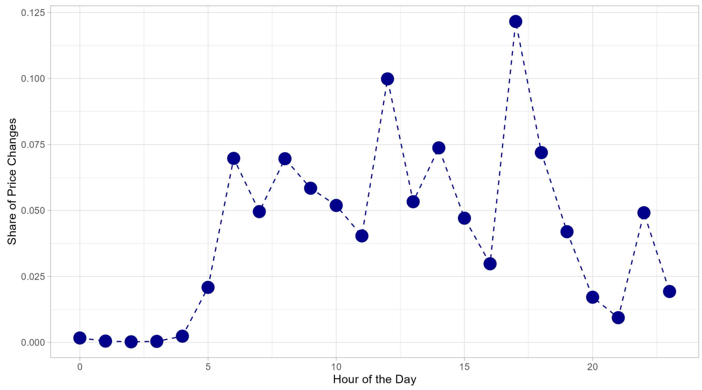


Figure 1.B.2: Timing of Price Changes

Note: This plot gives the timing of price changes of stations in the sample.

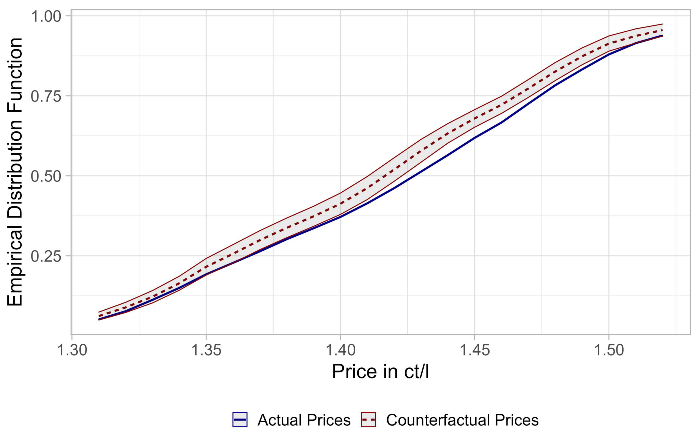


Figure 1.B.3: Distributional Effects of the Policy: Counterfactual Price Distribution - Midnight

Note: This plot gives the empirical distribution function of observed midnight post-lifting prices in Baden-Wuerttemberg (blue) and the counterfactual distribution for a scenario without policy lifting (red). The counterfactual distribution comes from distribution regression in the style of Chernozhukov et al. (2013) in one Eurocent/l steps. We provide 95% confidence intervals. Standard errors are clustered at the county level. The distributions are trimmed at the 5th and 95th percentile.

Appendix C: Opening Hours - Robustness Check

In our main analysis, we use price changes as a measure of stations' nightly activity. While price changes are only meaningful for stations, which are open at a certain time of the day, price changes might not perfectly reflect opening hours. For example, stations, which do not change prices at night, will not be identified as open stations then. To test the robustness of our findings above, we, therefore, provide further results on opening hours. Subsequently, we discuss the data we use and the analysis we conduct.

Data. We collect additional data on station-level opening hours from historic webpages of the price comparison website `clever-tanken.de` through the internet archive `web.archive.org`. The internet archive saves historic webpages erratically and inconsistently across webpages. But it allows us to extract historic opening hours from the gasoline stations' pages on `clever-tanken.de`. Each gasoline station has its own webpage on this domain, which is regularly updated in response to changing prices or opening hours. This website holds up-to-date opening hours since the data is provided by either consumers or the Federal Cartel Office. We can match the opening hours to stations in our dataset based on the name, address, and brand of the stations in the URL code of the archived webpages.
As the stations' sites are archived unregularly, we cannot retrieve opening hours for each station in the relevant time period. Hence, the panel is unbalanced and might be prone to selection issues, which we will investigate later on. Overall, our sample ranges from mid-2016 to mid-2018 and consists of more than 3,500 stations ($\approx$ 25% of all stations) and 16,199 observations after subtracting highway stations and stations not able to increase opening hours. On average, each station's opening hours are observed four to five times in the sample. We construct an opening hour measure $1[\text{Active between 10pm and 5am}]_{st}$ which will turn one if a station s opens at least once per week during the prohibition period 10pm and 5am on the scraped website from date t. The outcome definition follows the definition from the main analysis in Section 1.5 where a station was active when setting at least one price between 10pm and 5am per week. Moreover, we add two alternative outcomes to test the robustness of our results. First, the logged number of days per week a station opens during the nightly prohibition time. Second, a dummy for a station opening on at least five days per week during the nightly prohibition time.

Results. Table 1.C.1 gives the simple difference-in-differences effects of the treatment on the likelihood to open in the prohibition period for a sample of all stations and two subsamples of stations with small or large shop assortment (classifications as in the paper above). The regression follows the approach in equation (1.4). Similar to Figure 1.4 in the paper, we find opening at night to become more likely in response to the treatment. The treatment increases the likelihood to open at night by 1.5 percentage points (s. column (1)). In line with Figure 1.5, the effect is especially driven by shops with a small assortment (see columns (2) and (3) of Table 1.C.1). For the two alternative outcomes, a significant opening hour effect is also only evident for shops with a small assortment.

Table 1.C.1: Robustness Check: Opening Hours

	1[Active between 10pm and 5am]			ln(# Days/Week Open betw. 10pm and 5am)			1[Active > 5 Days/Week betw. 10pm and 5am)]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BW×Post	0.0154*	0.0118	0.0173**	0.0271	0.0168	0.0335*	0.0069	−0.0000	0.0118***
	(0.0091)	(0.0194)	(0.0087)	(0.0223)	(0.0551)	(0.0186)	(0.0106)	(0.0241)	(0.0044)
Sample	All	Large Assortment	Small Assortment	All	Large Assortment	Small Assortment	All	Large Assortment	Small Assortment
Station FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	16,199	6,331	9,868	16,199	6,331	9,868	16,199	6,331	9,868
Adjusted R^2	0.982	0.983	0.982	0.985	0.984	0.985	0.981	0.978	0.982

Note: *p<0.1; **p<0.05; ***p<0.01. All results are based on OLS regressions with standard errors clustered at the county level. The regression setup follows a simple DID. To include zero value observations in a logged transformation, we use the inverse hyperbolic sine transformation.

To provide support that these effects can be interpreted as causal effects of the policy lifting, Figures 1.C.1 and 1.C.2 show rather flat pre-trends for the pooled effect and the subsamples. We also show that this effect is robust to extending the sample period by six additional months after the policy lifting. Figure 1.C.3 shows that the positive effect for small assortment stations is persistent over time.

Note that the effects found are smaller than those on the likelihood of price changes in Figures 1.4 and 1.5. A potential reason for this is that the treatment effect on the likelihood of price changes there is upward biased. For example, stations which opened at night before the policy lifting but only start to change prices at night after the policy, are misleadingly detected as stations that extend opening hours. A reason might be that a station is located near a station, which extends opening hours in response to the policy. The incumbent station might then react by changing prices as well without extending opening hours. That is, opening hour effects likely spill over to the price setting of others, so that an upward bias in the regressions on price changes is likely.

Sample Selection. To understand in how far our sample is representative for the overall population of stations, we compare the characteristics of the sampled stations to the overall population. Table 1.C.2 provides evidence that the sample of historic prices includes stations from more strongly contested markets with slightly lower prices right before the treatment. In strongly contested markets, we did not find a price effect in response to the policy (s. Section 1.5). If such markets react less strongly to the policy, our sample might underestimate the opening hour effect for the general sample. On the other hand, sampled stations are located in counties with an on average higher share of youths, which may increase the reaction to the policy. Hence, ex-ante it is not clear that the selected sample induces an over- or

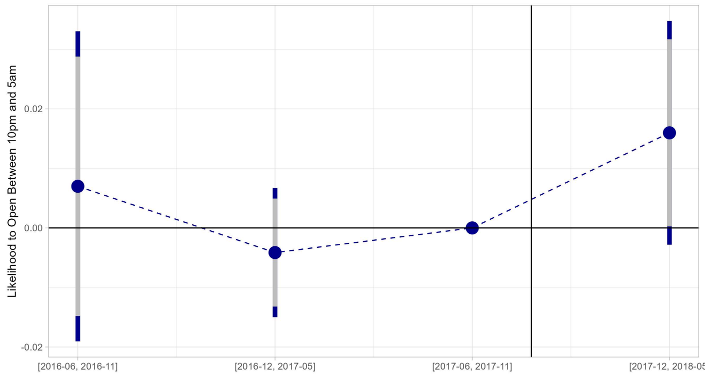


Figure 1.C.1: Robustness Check: Opening Hours

Note: This plot gives dynamic estimates of the leads and lags for the pooled sample. The outcomes variable is a dummy and turns one if a station opens between 10pm and 5am at least once a week. Standard errors are clustered at the county level. The exact timing of the beginning and end of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients from a linear probability model.

underestimation of the population effect. Importantly, the sample is similar in the share of premium stations - so stations with a large shop assortment.

Table 1.C.2: Representativeness of Historic Opening Hours Data

	Scraped Data	Full Sample	Δ (p-value)
1[Baden-Wuerttemberg]	0.148	0.130	0.12
E5 Gasoline Price (Day)	1.361	1.366	0.00***
# Competitors 0.5km Radius (Day)	0.424	0.402	0.05*
# Competitors 0.5km Radius (Night)	0.163	0.164	0.98
# Competitors 1km Radius (Day)	1.134	0.986	0.00***
# Competitors 1km Radius (Night)	0.447	0.402	0.00***
Premium Station	0.477	0.494	0.05*
Oligopolistic Station	0.459	0.434	0.02**
Share of Youths (18-25 y.o., County Level)	0.091	0.087	0.00***

Note: *p<0.1; **p<0.05; ***p<0.01. P-values of differences come from regressions of the outcome on a constant and a dummy for stations that are part of the scraped dataset. Standard errors are clustered at the county level. Prices are average prices from the last pre-treatment week for such stations which operated in this week.

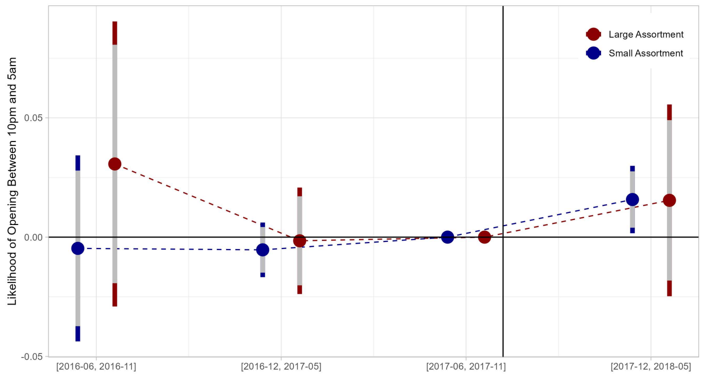


Figure 1.C.2: Robustness Check: Opening Hours - Heterogeneity

Note: This plot gives dynamic estimates of the leads and lags for two subsamples of stations with heterogeneous store assortment. The outcomes variable is a dummy and turns one if a station opens between 10pm and 5am at least once a week. Standard errors are clustered at the county level. The exact timing of the beginning and ending of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients from a linear probability model.

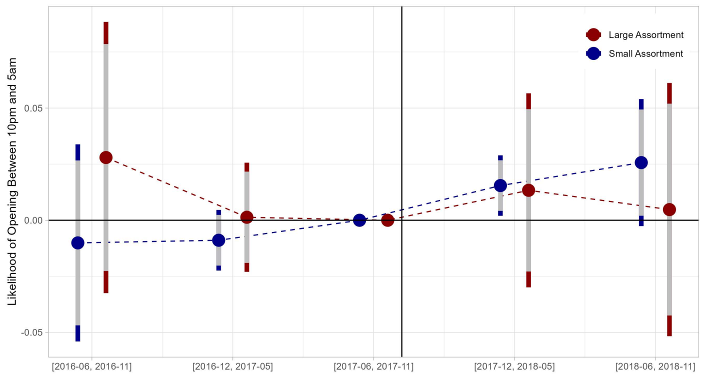


Figure 1.C.3: Robustness Check: Opening Hours - Extended Sample Period

Note: This plot gives dynamic estimates of the leads and lags for the pooled sample and two subsamples of stations with heterogeneous store assortment. The outcomes variable is a dummy and turns one if a station opens between 10pm and 5am at least once a week. Standard errors are clustered at the county level. The exact timing of the beginning and ending of the prohibition is indicated by the black vertical line. We provide 90 and 95% confidence intervals for all coefficients from a linear probability model.

Appendix D: Effect on Maximum Willingness to Pay

This section sets up and tests hypotheses of how a policy-induced change in the share of informed consumers affects the theoretical, market-level price distribution. In a second step, we test in a difference-in-differences setup whether the effects in the data are better explained by a change in informed consumers or an increasing valuation of gasoline.

Information - in this setting - means the share of consumers who actually are aware of the gasoline price. The policy might induce that a higher share of consumers actually visits gasoline stations expecting to only get alcohol. Also, more consumers might not care about the gasoline price after the policy lifting. Then, the policy would increase the share of consumers, who buy at a random price.

We build on the canonical search model by Varian (1980)[17]. His model sets up a market of N symmetric firms, which each sell a homogenous product to a unit mass of consumers. All consumers have a willingness to pay of v. In our case, the product is gasoline, which arguably is homogenous (Martin (2023), Montag et al. (2023), Pennerstorfer et al. (2020)) and demand is quite inelastic. A share η of consumers know all prices in the market and hence buy at the lowest price. $1-\eta$ consumers do not observe prices and buy at a random station. In our case, the policy lifting could induce that more consumers do not expect to buy gasoline and hence buy at a random station. This is equivalent to a lower η.

Following Varian (1980), the unique equilibrium is given by the price distribution

$$F(p) = 1 - \left(\frac{1-\eta}{N\eta}\frac{v-p}{p-c}\right)^{\frac{1}{N-1}},$$

where c are marginal costs and the support of $F(p)$ is given by $[\underline{p}, \bar{p}] = [c + \frac{v-c}{1+\frac{N\eta}{1-\eta}}, v]$.

One can easily see that the upper bound of the support of $F(p)$ is independent of the share of informed consumers. Changes in the upper bound can only be rationalised through changes in consumers' valuation for gasoline.

We test empirically whether the upper bound of prices $\bar{p}$ is affected by the treatment. As argued in Wildenbeest (2011), the maximum price observed in a market is a consistent estimate of the upper bound of $F(p)$. For each market and month, we get the estimate as the maximum price observed[18]. We abstract from markets with only one firm as the model's predictions do not hold for monopolies. Markets are defined at the station level by drawing circles around each station. We apply a 0.5km, 1km, and 2km radius.

We, then, run the following regression

$$\hat{\bar{p}}_{mt} = \alpha_m + \lambda_t + \theta(BW_m \times Post_t) + \epsilon_{mt},$$

[17] Other papers such as Pennerstorfer et al. (2020) model consumer information in gasoline markets similarly.

[18] Note that the estimate heavily varies with the oil price. Nevertheless, we are only interested in the difference and comparison between markets in Baden-Wuerttemberg and other federal states. Hence, the level of the estimate is less relevant as we account for month fixed effects in the difference-in-differences regression later on.

Table 1.D.1: Effect on Monthly, Market-Level Willingness to Pay

	$\hat{\bar{p}}$		
	(1)	(2)	(3)
BW×Post	0.0058*	0.0049**	0.0070***
	(0.0032)	(0.0023)	(0.0020)
Market	0.5km Radius	1km Radius	2km Radius
Market FE	✓	✓	✓
Month FE	✓	✓	✓
Observations	11,066	22,986	40,709
Adjusted R^2	0.871	0.771	0.767

Note: *p<0.1; **p<0.05; ***p<0.01. As the model only predicts outcomes for markets with at least two competitors, we drop markets with only one station. Maximum prices based on daily, station-level midnight prices. We provide 90 and 95% confidence intervals. Standard errors are clustered at the market level. Related literature uses similar market radii for market delineation such as 2 miles linear distance (Chandra and Tappata (2011)) or 2 miles driving distance (Pennerstorfer et al. (2020)). December 2017 is the first month classified as post-lifting.

where $\hat{\bar{p}}_{mt}$ is the estimate for $\bar{p}$ in market m and month t. Table 1.D.1 shows the treatment effect for different market definitions. Results show that that $\hat{\bar{p}}$ increases with the treatment. This is not in line with a change in the share of informed consumers but with an increasing valuation for gasoline. This supports the demand expansion mechanism described in the main part of the paper.

Appendix E: Dynamic Rank Reversal Test

In a homogenous product market, price rank reversals of two gasoline stations can hardly be explained without considering consumer frictions. For example, input price changes over time do not cause prices of one station to increase more than those of others.
Chandra and Tappata (2011) propose a test for consumer frictions following this intuition. For each station couple c (station A and station B), one calculates the share of days, on which the usually cheaper station A is the more expensive one. Formally, this is given by:

$$rr_{ct} = \frac{1}{N_{ct}} \sum_{\tau=1}^{N_{ct}} 1[p_{A\tau} > p_{B\tau}]$$

where rr_{ct} gives the rank reversal measure for couple c for period t. N_{ct} is the number of days both stations report a price and $p_{A\tau}$ and $p_{B\tau}$ are prices at midnight for stations A and B on day τ. We calculate the rank reversal measure for each couple twice - once for all observations before and once for all observations after the policy lifting. Only data from dates on which both stations operate is used. We use data from midnight prices. As prices hardly change during nighthours (s. Figure 1.B.2), this analysis likely holds for all other points in time during the nightly prohibition.

Table 1.E.1: Effect of Policy Lifting on Likelihood of Rank Reversals

	rr_{ct}			
	(1)	(2)	(3)	(4)
$1[1 \geq$ Distance in km $\leq 1.5]$	-0.0109^{**}			
	(0.0043)			
$1[0.5 \geq$ Distance in km $\leq 1]$	-0.0108^{**}			
	(0.0047)			
$1[0.15 \geq$ Distance in km $\leq 0.5]$	-0.0159^{***}			
	(0.0060)			
$1[$Distance in km $\leq 0.15]$	-0.0260^{***}			
	(0.0083)			
BW×Post		-0.0040	0.0101	0.0090
		(0.0194)	(0.0122)	(0.0084)
Couple Distance	$\leq$ 2km	$\leq$ 0.5km	$\leq$ 1 km	$\leq$ 2km
Couple FE	×	✓	✓	✓
Observations	7,073	991	2,513	7,073
Adjusted R^2	0.002	0.408	0.359	0.345

Note: *p<0.1; **p<0.05; ***p<0.01. Rank reversal measures based on daily prices at midnight. Standard errors are clustered at the station couple level. Post dummy included in the regressions.

We then run the following regressions for a subsample of couples with a maximum linear distance between the two stations of 1km or 2km:

$$rr_{ct} = \lambda_c + \gamma Post_t + \beta BW_c \times Post_t + \epsilon_{ct}$$

where β gives the change in rank reversals related to the policy lifting. A couple is considered to belong to Baden-Wuerttemberg ($BW_c = 1$) if both stations are located in Baden-Wuerttemberg but our results also hold when a couple is also treated in the case that only one station lies in Baden-Wuerttemberg.
Table 1.E.1 shows two results: First, column (1) shows that frictions decrease for a lower distance between stations which is in line with findings in the literature (Chandra and Tappata (2011), Martin (2023), Pennerstorfer et al. (2020)). Second, we show that the policy does not affect rank reversals in columns (2) to (4). Hence, there is no indication for a change in consumer information caused by the policy.

Appendix F: Correlation of Prices

In this section of the appendix, we show that prices of neighboring stations do not comove more or less in response to the policy lifting.

Demand could have become less elastic in Baden-Wuerttemberg after the policy lifting as more consumers visit the gasoline station for alcohol and, hence, might care less about the gasoline price. If this was the case, prices of competitors would become less important (less elastic cross-price elasticity). At the extreme, demand could be sufficiently inelastic so that stations become quasi-monopolists. Then, neighboring stations' prices will not be strategic responses.

To examine whether neighboring stations' price comovement is affected by the policy lifting, we setup the following difference-in-differences regression equation:

$$corr_{ct} = \lambda_c + \gamma Post_t + \beta BW_c \times Post_t + \epsilon_{ct}$$

where $corr_{ct}$ is the correlation between midnight prices of neighboring stations for two time periods t - before and after the treatment - separately. We later on use the correlation of actual prices as well as prices residualized from state-date specific price effects. The latter accounts for correlation purely caused by input price fluctuations. β gives the treatment effect.

Table 1.F.1 gives the estimation results. There is no robust evidence for a change in the correlation of neighboring stations' prices. This supports our claim that the price effect is not caused by less elastic demand in response to the policy. This would have likely resulted in a lower correlation of prices after the policy lifting in Baden-Wuerttemberg.

Table 1.F.1: Effect of Policy Lifting on Neighboring Stations' Price Correlation

	$corr_{ct}$					
	Actual Prices			Residualized Prices		
	(1)	(2)	(3)	(4)	(5)	(6)
BW×Post	0.0341	0.0157	0.0288**	−0.0343	−0.0200	−0.0114
	(0.0464)	(0.0240)	(0.0145)	(0.0524)	(0.0367)	(0.0217)
Couple Distance	≤ 0.5km	≤ 1 km	≤ 2km	≤ 0.5km	≤ 1 km	≤ 2km
Couple FE	✓	✓	✓	✓	✓	✓
Observations	957	2,476	7,017	990	2,510	7,062
Adjusted R^2	0.449	0.496	0.521	0.652	0.622	0.623

Note: *p<0.1; **p<0.05; ***p<0.01. Correlations calculated based on daily, station-level midnight prices. Standard errors are clustered at the station couple level. Post dummy included in the regressions.

Bibliography

Arcidiacono, P., Ellickson, P. B., Mela, C. F. and Singleton, J. D. (2020), 'The Competitive Effect of Entry: Evidence from Supercenter Expansion', *American Economic Journal: Applied Economics* **12**(3), 175–206.

Armstrong, M. and Vickers, J. (2012), 'Consumer Protection and Contingent Charges', *Journal of Economic Literature* **50**(2), 477–493.

Assad, S., Clark, R., Ershov, D. and Xu, L. (2021), 'Algorithmic Pricing an Competition: Empirical Evidence from the German Retail Gasoline Market', *cesifo Working Paper No. 8521* .

Baueml, M., Marcus, J. and Siedler, T. (2023), 'Health Effects of a Ban on Late-Night Alcohol Sales', *Health Economics* **32**(1), 65–89.

Baumann, F., Buchwald, A., Friehe, T., Hottenrott, H. and Mechtel, M. (2019), 'The Effect of a Ban on Late-Night Off-Premise Alcohol Sales on Violent Crime: Evidence from Germany', *International Review of Law and Economics* **60**, 105850.

Bundesregierung (2018), 'Bericht über die Ergebnisse der Arbeit der Markttransparenzstelle für Kraftstoffe und die hieraus gewonnenen Erfahrungen', *Drucksache 19/3693* .

Chandra, A. and Tappata, M. (2011), 'Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets', *RAND Journal of Economics* **42**(4), 681–704.

Chernozhukov, V., Fernandez-Val, I. and Melly, B. (2013), 'Inference on Counterfactual Distributions', *Econometrica* **81**(6), 2205–2268.

Doyle, J. J. J., Muehlegger, E. and Samphantharak, K. (2010), 'Edgeworth Cycles Revisited', *Energy Economics* **32**(3), 651–660.

FAZ (2015), 'Tankstellen ärgern sich über Benzinpreise', *Frankfurter Allgemeine Zeitung* .

Federal Cartel Office (2011), *Sektorenuntersuchung Kraftstoffe*, Bonn.

Federal Cartel Office (2018), *Markttransparenzstelle für Kraftstoffe (MTS-K): Jahresbericht 2017*, Bonn.

Federal Cartel Office (2019), *Markttransparenzstelle für Kraftstoffe (MTS-K): Jahresbericht 2018*, Bonn.

Fischer, K., Martin, S. and Schmidt-Dengler, P. (2023), 'The Heterogeneous Effects of Entry on Prices', *CESifo Working Paper 10566* .

Fuest, C., Peichl, A. and Siegloch, S. (2018), 'Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany', *American Economic Review* **108**(2), 393–418.

Gabaix, X. and Laibson, D. (2006), 'Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets', *The Quarterly Journal of Economics* **121**(2), 505–540.

Garthwaite, C., Gross, T. and Notowidigdo, M. (2014), 'Public Health Insurance, Labor Supply and Employment Lock', *Quarterly Journal of Economics* **129**, 653–696.

Goolsbee, A. and Syverson, C. (2008), 'How Do Incumbents Respond to the Threat of Entry?', *Quarterly Journal of Economics* **123**(4), 1611–1633.

Haucap, J., Heimeshoff, U. and Siekmann, M. (2017*a*), 'Fuel Prices and Station Heterogeneity on Retail Gasoline Markets', *Energy Journal* **38**(6), 81–103.

Haucap, J., Heimeshoff, U. and Siekmann, M. (2017*b*), 'Selling Gasoline as a By-Product: The Impact of Market Structure on Local Prices', *DICE Discussion Paper 240* .

Heidhues, P., Koszegi, B. and Murooka, T. (2017), 'Inferior Products and Profitable Deception', *Review of Economic Studies* **84**, 323–356.

Isphording, I. E., Lipfert, M. and Pestel, N. (2021), 'Does Re-Opening School Contribute to the Spread of SARS-CoV-2? Evidence from Staggered Summer Breaks in Germany', *Journal of Public Economics* **198**, 104426.

Ivanov, A. (2019), 'Lohnt es sich noch, eine Tankstelle zu betreiben?', *Handelsblatt* .

Korff, A. (2021), 'Competition on the Fast Lane: The Price Structure of Homogeneous Retail Gasoline Stations', *DICE Discussion Paper 359* .

Lal, R. and Matutes, C. (1994), 'Retail Pricing and Advertising Strategies', *Journal of Business* **67**(3), 345–370.

Landtag von Baden-Wuerttemberg (2020), 'Folgen der Aufhebung des nächtlichen Alkoholverkaufsverbots und Umsetzung der Einführung von Alkoholkonsumverboten in Baden-Würrtemberg'.

Lichter, A. and Schiprowski, A. (2021), 'Benefit Duration, Job Search Behavior and Re-Employment', *Journal of Public Economics* **193**, 104326.

Marcus, J. and Siedler, T. (2015), 'Reducing Binge Alcohol? The Effect of a Ban in Late-Night Off-Premise Alcohol Sales on Alcohol-Related Hospital Stays in Germany', *Journal of Public Economics* **123**, 55–77.

Martin, S. (2023), 'Market Transparency and Consumer Search - Evidence from the German Retail Gasoline Market', *RAND Journal of Economics* **55**(4), 573–602.

Matsa, D. A. (2011), 'Competition and Product Quality in the Supermarket Industry', *Quarterly Journal of Economics* **126**(3), 1539–1591.

Mayer, N. (2017), 'Land kippt Verkaufsverbot erst 2018', *Stuttgarter Nachrichten* .

Montag, F., Mamrak, R., Sagimuldina, A. and Schnitzer, M. (2023), 'Imperfect Price Information, Market Power, and Tax Pass-Through'. Working Paper.

Nicolai, B. (2021), 'Alkohol, Izgaretten, Klopapier - was die Tanke jetzt über unser Leben verrät', *Welt* .

Ning, X. and Haining, R. (2003), 'Spatial Pricing in Interdependent Markets: A Case Study of Petrol Retailing in Sheffield', *Environment and Planning A: Economy and Space* **35**(12), 2131–2159.

Noel, M. D. (2007), 'Edgeworth Price Cycles, Cost-based Pricing and Sticky Pricing in Retail Gasoline Markts', *Review of Economics and Statistics* **89**(2), 324–334.

NTV (2015), 'Tankstellen-Pächter zahlen für Jo-Jo-Preise', *NTV.de* .

Olden, A. and Moen, J. (2022), 'The Triple Difference Estimator', *Econometrics Journal* **25**(3), 531–553.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C. and Yontcheva, B. (2020), 'Information and Price Dispersion: Theory and Evidence', *International Economic Review* **61**(2), 871–899.

Scope Ratings (2018), *Branchenstudie Tankstellenmarkt Deutschland 2017*.

Scope Ratings (2019), *Branchenstudie Tankstellenmarkt Deutschland 2018*.

Siekmann, M. (2017), 'Characteristics, Causes, and Price Effects: Empirical Evidence of Intraday Edgeworth Cycles', *DICE Discussion Paper 252* .

Statistisches Landesamt Baden-Württemberg (2022), 'Kraftfahrzeugbestand und Kraftstoffverbrauch'. Accessed on 2022-10-13.
URL: *https://www.statistik-bw.de/Verkehr/KFZBelastung//LRt1506.jsp?path=Energie/Energiebilanz/*

Varian, H. R. (1980), 'A Model of Sales', *American Economic Review* **70**(4), 651–659.

Wang, Z. (2015), 'Supermarkets and Gasoline: An Empirical Study of Bundled Discounts', *Resources for the Future Discussion Paper 15-44* .

Wildenbeest, M. R. (2011), 'An Empirical Model of Search with Vertically Differentiated Products', *RAND Journal of Economics* **42**(4), 729–757.

Zimmerman, P. R. (2012), 'The Competitive Impact of Hypermarket Retailers on Gasoline Prices', *The Journal of Law and Economics* **55**(1), 27–41.

Chapter 2

The Heterogeneous Effects of Entry on Prices

Coauthor(s): Simon Martin, Philipp Schmidt-Dengler (both University of Vienna)

Abstract: We study the effect of entry on the price distribution in the German retail gasoline market. Exploiting several hundred entries over five years in an event study design, we find that entry causes a persistent first-order stochastic shift in the price distribution for at least three years after entry. Prices at the top of the distribution change moderately or not at all, but prices at the left tail decrease by up to 13% (9%) of stations' gross margins after entry within a 1 km (2 km) radius. The Value of Information (VOI) increases by 29% (15%) in 1 km (2 km) markets, suggesting larger savings for consumers with easy access to information.

2.1 Introduction

How changes in market structure affect competitive outcomes and consumer rents is a long-standing question in economics. The overall welfare effect of an increase in the number of firms, i.e. entry, in oligopoly settings is ambiguous as entry involves fixed costs but generally lowers prices and improves product quality (Mankiw and Whinston, 1986). Therefore consumers, on average, benefit from entry.

However, when not all consumers are equally informed about prices in the market, their gains of entry are potentially heterogeneous. To illustrate this idea, suppose that there are two types of consumers in a homogeneous product market (Varian, 1980). *Uninformed* consumers do not compare prices and simply purchase from a firm at random. For these consumers, the expected transaction price equals the average price in the market. *Informed* consumers, in contrast, observe all prices. The average transaction price of such a consumer corresponds to the expected *minimum* price in the market instead. In the presence of price dispersion, the effect of entry for these two groups depends on how entry affects the price distribution.[1]

Price dispersion is ubiquitous in markets for physically homogeneous goods to an extent that cannot be explained by differences in location, cost, or services. Thus, it is important to understand how entry affects the distribution of prices and, subsequently, how consumers holding different information about prices are heterogeneously affected.[2]

In this paper, we give an empirical answer. We investigate how information about prices mediates the benefits of entry to consumers in the German retail gasoline market. Using real-time price data for the universe of German gasoline stations over five years and exploiting several hundred time- and geo-coded station entries, we identify the effect of entry on the entire price distribution.

The retail gasoline market provides an ideal setting to tackle the question at hand. First, despite an arguably fairly homogeneous product, gasoline markets exhibit substantial price dispersion. In our sample, the average market-level price range is 2.7 ct/l (Eurocents per liter), corresponding to almost 40% of an average station's margin. Given that the average German household spends more than 1,000 Euro p.a. on fuel (Federal Ministry for Economic Affairs and Energy, 2021), distributional implications across heterogeneously informed consumers matter. Second, stations primarily compete on the price dimension where menu costs are effectively zero. Third, the retail gasoline industry consists of numerous small markets, in which entry by one more station is likely to have non-negligible price effects. The spatial nature of competition in combination with a lack of nationwide pricing allows us to find suitable control markets. Last, the several hundreds of entries into narrow, isolated markets enable us to study heterogeneous effects of entry depending on station and market characteristics.

[1] In the model of sales introduced by Varian (1980), average prices even rise upon an exogenous increase in the number of firms from N to $N + 1$, so that uninformed consumers are hurt by entry and only informed consumers benefit. Lach and Moraga-González (2017) show that this result does not extend to richer heterogeneity in consumer information.

[2] As consumer information and search behavior are strongly correlated with market-specific socio-economic variables such as income (Byrne and Martin, 2021, Nishida and Remer, 2018), heterogeneous effects along the information distribution imply different welfare effects across socio-economic backgrounds.

Exploiting the staggered nature of entry events in a difference-in-differences event study design, we first quantify the effect of entry on incumbent stations' average prices as a benchmark. Following Arcidiacono et al. (2020), the entry effect can be identified under the assumption that the precise timing of entry is exogenous from the incumbent's perspective. Leveraging the detailed within-market variation in competition, we estimate that average prices decrease by around 7% of a station's margin due to entry. This effect is persistent for at least three years post-entry.

Next, we explore the effect of entry on the price distribution at different quantiles by implementing the recentered influence function (RIF) regression approach of Firpo et al. (2009). To infer the counterfactual price distribution we employ the distribution regression approach by Chernozhukov et al. (2013). Entry reduces incumbents' prices at all quantiles of the price distribution. The post-entry price distribution is first-order stochastically dominated by the counterfactual distribution of prices absent entry. Hence, all consumers benefit, in expectation, from an increase in competition. However, entry affects the price distribution asymmetrically: The price decrease from entry is most pronounced at the left tail of the distribution where prices decrease by around 1 ct/l or 13% of stations' gross margins. At the top of the distribution, price changes due to entry are not significantly different from zero. As entry might additionally facilitate price comparisons due to more stations in the consumer's direct vicinity, our estimates are a conservative estimate of the gains from entry.

When consumers are differentially informed (i.e., they observe a different number of price draws before making a purchase decision), their expected *transaction prices* differ, even when they face the same *posted prices*. To quantify this heterogenous effect, we focus on the value of information (VOI): the difference between the average and the minimum posted prices in a market. This metric captures the gains of being fully informed about prices relative to observing only one price at random. We show that entry increases VOI by 29%. Hence, better-informed consumers benefit considerably more from entry.

The estimated effects are robust to a variety of sensitivity checks. We show that entry persistently increases the number of firms in the treated markets. Entry neither fosters nor anticipates exit, i.e. it does not take place in response to exit. Using hourly traffic data from almost 2,000 traffic counters in Germany, we also show that entry does not take place in response to a demand shift (Dubois et al., 2015), i.e. an increase in traffic in the market. Our estimates further show no effect of entry on incumbents' opening hours. This suggests that incumbents' primary response to entry is price adjustment. Beyond that, we show that entry also does not affect consumer information frictions in the market. Lastly, our findings are robust to various specifications, alternative region-time-specific controls, different estimation methods, and the time of the day analyzed. Finally, we carry out the analogous exercise to examine the effect of exit. The results on exit mirror the effects found for entry: Exit increases prices, and this effect is more pronounced at the lower end of the distribution.

The contributions of this paper are multifold. First, by causally uncovering the heterogeneous consumer gains from entry and competition, our paper informs policymakers about a new dimension in the welfare effects of competition policy. For example, merger control and entry regulation (indirectly) affect the number of competitors in a market. Our paper sheds light on the fact that the average price effects of policy measures can strongly differ from price effects for specific consumer groups.

We further contribute to debates on whether and how policymakers should provide information to consumers. In many industries, policymakers have launched price comparison tools, e.g. for financial products or energy contracts. Also in the gasoline markets of, for example, Germany, France, Austria, or Western Australia, price comparison tools for consumers were introduced to reduce search costs. While most academic research focuses on the question of whether transparency policies foster firm coordination more than consumers' ability to find the lowest price (Ivaldi et al., 2003, Kühn and Vives, 1995, Luco, 2019, Rossi and Chintaguanta, 2016), our paper unveils a positive effect of such transparency policies: Marginal consumers who become better informed due to such tools benefit more from an increase in competition ceteris paribus. Hence, transparency policies have different effects on (un)informed consumers depending on the competition intensity in a market.

Second, we contribute to research studying the effect of firm entry on prices. The largest body of research analyses retail grocery and the airline industry. Basker (2005) and Basker and Noel (2009) show that Walmart's entry reduces competitor prices as well as market-level average prices. Arcidiacono et al. (2020) find that Walmart's entry substantially reduces revenues by incumbent grocery stores, but they do not identify a significant price response. Bauner and Wang (2019) and Oschmann (2022) show that incumbents instead react to Walmart or Costco entry by reducing product variety and adapting their pricing strategies to more fluctuating prices. Atkin et al. (2018) and Busso and Galiani (2019) study retail markets in Mexico and find price effects of entry in the range of 2-6%. Bernardo (2018), Cardoso et al. (2022) and Davis et al. (2022) investigating gasoline markets in Spain, Brazil, and Mexico, respectively, similarly find a negative average price effect of entry like us, but do not analyze distributional effects.

For the airline industry, Whinston and Collins (1992) show that incumbent airlines lower prices on the entered routes. Goolsbee and Syverson (2008) establish that incumbents lower prices already in response to the mere threat of entry by a discount carrier. Prince and Simon (2015) show that this response is accompanied by a reduction in service quality as measured by the on-time performance of incumbents. Reiffen and Ward (2005) employ the timing of FDA approval as exogenous in the pharmaceutical industry to establish how generic entry competes away price-cost margins. We complement this literature by looking at a novel information channel that shapes entry effects on consumers. Additionally, we provide new evidence on the average price effects of entry based on hundreds of entry events in narrowly defined markets in which regional and national pricing are less concerning.

Third, we add to the empirical literature on the interaction of competition, informational frictions, and price dispersion, especially in gasoline markets, such as Barron et al. (2004), Chandra and Tappata (2011), Lach and Moraga-González (2017), Lewis (2008), and Pennerstorfer et al. (2020). Rather than relying on cross-sectional variation in the number of competitors between markets, our event study design exploits within-firm and within-market price variation and isolates the causal effect of entry. Moreover, the setting enables us to explicitly study how consumers are differently affected by changes in competition, depending on how informed they are about prices. While Barron et al. (2004) and Lewis (2008) find a negative relation between competition and price dispersion, we find causal evidence for a positive relation, which is in line with Pennerstorfer et al. (2020).

Finally, our paper complements the few papers which explicitly examine the effect of competition on the whole price distribution. Allen et al. (2014) study a bank merger and its heterogeneous effect on mortgage prices across local markets. In contrast to us, they exploit the ownership structure in the market, but not changes in the number of outlets. Moreover, the entry of different types of stations into sub markets allows us to study how entrant and incumbent characteristics affect the change of the price distribution. Lach and Moraga-González (2017) explore how the number of competitors in local gasoline markets affects the price competition in cross-sectional analysis between markets. Similar to us, they find that competition shifts prices in a first-order stochastic manner. Contrary to us, in their paper prices in the right tail of the distribution decrease the most with competition.

The remainder of the paper proceeds as follows: In Section 2.2, we discuss the institutional setting of our research as well as the data we use. We then proceed with discussing the empirical strategy in Section 2.3. In Section 2.4, we present our results on the effect of entry on the price distribution, we discuss channels that give rise to these results, and we present several robustness checks. Finally, we conclude in Section 2.5.

2.2 Industry Background and Data

We study the effect of entries into the German gasoline market from 2015 to 2020 by combining data from various sources.[3]

We obtain the universe of prices and detailed information for all gasoline stations (brand affiliation, address, exact location) over the entire sample period from Tankerkönig (2025).[4] The market is dominated by five major brands: Aral (BP), Shell, Total, Esso, and JET, which together operate almost half of all stations in Germany (see Figure 2.B.1 in the Appendix for a detailed overview). Around 25% of stations are not affiliated with any brand and are therefore considered independent. Since early 2017, algorithmic pricing software is increasingly utilized (Assad et al., 2024).[5]

Gasoline stations on highways are arguably competing for a different set of consumers (Bun-

[3]Our sample ends in early 2020, the beginning of the COVID-19 pandemic. We drop the period due to substantial demand shocks as well as a temporary VAT reduction in 2020. See Montag, Mamrak, Sagimuldina and Schnitzer (2023) and Fischer et al. (2024) for studies on this VAT reduction.

[4]The same data source is also used e.g. in Montag, Sagimuldina and Winter (2023), Montag, Mamrak, Sagimuldina and Schnitzer (2023) and Martin (2024). This database uses real-time data provided by the German Market Transparency Unit (MTU, in German Markttransparenzstelle für Kraftstoffe MTS-K), which is a sub-unit of the German competition authority (Bundeskartellamt). By regulation, gasoline stations are required to submit all price changes to the MTU 'instantanenously' since December 2013. The introduction of this regulation sufficiently predates our sample period and should therefore be inconsequential for our subsequent analysis.

[5]Although pricing software might effect the industry in many ways, we show in Section 4.5 the entry effects we are interested in in this paper seem orthogonal to the adoption of AI software.

deskartellamt, 2011, Martin, 2024, Montag, Mamrak, Sagimuldina and Schnitzer, 2023). We follow the literature and omit these stations from our analysis entirely.[6] Throughout, we focus on diesel prices at 5 pm, which is when most consumers fuel their cars (Deutscher Bundestag, 2018). Later we also provide results for E5 gasoline and other times of the day (see Section 2.4.7).

Table 2.1 provides summary statistics on the stations in our sample, 15,437 in total, where approximately 40% are open 24/7. Across all stations, the average posted price over the sample period is 117 ct/l. Prices at major brands (Total, Shell, Esso, Aral) tend to be higher, with an average brand premium of around 1 ct/l relative to unbranded stations, and even around 2 ct/l relative to budget stations such as JET stations.[7]

Gross margins account for around 6% of the retail price. Similar to Assad et al. (2024), we use daily, diesel wholesale price data for nine price regions in Germany as published in the Oil Market Report by Argus Media (2025) to calculate gross margins.[8] We match the closest, available regional wholesale price to each station on a daily basis. The average margin of slightly more than 7 ct/l fits industry survey evidence of approximately 10 ct/l (Scope Investor Services, 2021).

Using stations' address information and location data, we compute the number of competitors within a given radius (Pennerstorfer et al., 2020). The average station has 0.9 (2.5) competitors within a 1 km (2 km) radius, respectively. The distribution of market sizes across all market-date observations is illustrated in Figure 2.B.2 in the Appendix. According to a market delineation based on a 1 km (2 km) radius, 47% (25%) of markets are local monopolists, and 30% (21%) are duopolies respectively. Consequently, the majority of stations face local competition, even within fairly small radii.

In order to examine whether entry events are related to local changes in demand, we collect data proxying local fuel demand from two sources: First, we have traffic flow data from almost 2,000 traffic counters in Germany operated by Bundesanstalt für Straßen- und Verkehrswesen (2025). Second, we have access to very fine-grained annual population data at the 1x1 km-grid-level from the RWI-GEO-GRID dataset (Breidenbach and Eilers, 2018, RWI, 2023). Both data sources allow us to capture local trends within narrowly defined gasoline markets and also across municipality borders.

Finally, we complement our data with annual, county-level variables from the Statistical Offices of Germany and the Federal States. This data allows us to control for different local trends in GDP per capita, income, unemployment, population, vehicles, and commuting over

[6] We identify highway stations in a two-step procedure: First, we check stations on the website `https://www.raststaetten.de/alle-standorte` and `https://serways.de/standorte`, which together list all highway stations let out to tenants by *Tank & Rast*, the company in charge of all highway roadhouses (Bundeskartellamt, 2011), in a cross-section of February 2023. We get the exact coordinates of these roadhouses and their stations and calculate distances to stations in our dataset. Stations with a distance of below 400 metres to a roadhouse are coded as highway stations. In the second step, we manually check our results for errors. We examine the address list of all German stations to identify stations ID's which were not yet coded as highway stations or falsely coded as such. We classify almost 500 station IDs as highway gas stations and drop them.

[7] See Figure 2.D.1 in the Appendix for a more detailed distribution of brand premia.

[8] The gross margin is given by: $margin_{it} = price_{it}/(1 + VAT) - wholesaleprice_{it}$, where the wholesale price already includes the diesel excise tax.

Table 2.1: Station-Level Information

	Station Cross-Section			Station×Date Panel		
	N	Mean	S.D.	N	Mean	S.D.
Station Choices						
Price (ct/l)	15,437	117.08	3.118	25,515,289	116.83	10.274
TOTAL	849	117.70	2.810	1,507,106	117.43	10.297
SHELL	1,714	117.64	2.447	3,070,060	117.57	10.228
ESSO	1,069	117.11	2.306	1,920,473	117.13	10.398
ARAL	2,255	118.19	2.307	4,042,296	118.20	10.539
JET	692	115.84	2.500	1,198,634	115.55	10.033
Others	8,858	116.72	3.448	13,776,720	116.27	10.148
Gross Margin (ct/l)	15,437	7.449	1.639	25,515,289	7.401	3.085
1[Open 24/7]	14,835	0.394	0.483	5,996,285	0.382	0.486
Station Characteristics						
# Competitors 1 km Radius	15,437	0.869	1.079	25,515,289	0.885	1.093
# Competitors 2 km Radius	15,437	2.500	2.599	25,515,289	2.574	2.640
1[Big Four]	15,437	0.381	0.486	25,515,289	0.413	0.492
1[JET]	15,437	0.047	0.211	25,515,289	0.048	0.215
1[Entrant]	15,437	0.046	0.210	25,515,289	0.024	0.154
1[Incumbent - 1 km Radius]	15,437	0.035	0.183	25,515,289	0.034	0.181
1[Incumbent - 2 km Radius]	15,437	0.086	0.280	25,515,289	0.086	0.280

Note: To obtain station-level, cross-sectional variables, we averaged over the respective observations at the station level. Using daily data from early 2015 to early 2020, we observe approximately 1,850 dates in our sample.

time (Statistical Offices of Germany and the Federal States, 2025*a*,*b*,*c*,*d*,*e*,*f*). We also obtained data on the number of vehicles and commuters at the county level.

2.2.1 Price Dispersion

An important characteristic of retail gasoline markets is cross-sectional price dispersion, which is substantial even though gasoline is a fairly homogeneous good. We report several measures of price dispersion at the market level, namely the sample standard deviation of prices ($S.D.$), the range of prices (*Range*), and the value of information (VOI), i.e., the difference between the *average price* and the *minimum price*, in Table 2.2.
We define markets as circles of 1 km and 2 km radius around each station.[9] Within a 2 km radius, the average price range is around 2.7 ct/l, i.e., 2% of the sample average price or almost 40% of stations' average gross margin (Scope Investor Services, 2021).[10] The average value of information is 1.2 ct/l. This implies that consumers could gain considerably, depending

[9]Similar approaches and distances are used in Chandra and Tappata (2011), Moraga-González et al. (2017), Cabral et al. (2019) and Pennerstorfer et al. (2020). In our analyses, we drop markets where entrants are the focal station, as we naturally do not observe these markets before entry events.

[10]In some markets, dispersion is 0 ct/l (see values at first decile). This is due to the fact that most markets in our sample are duopolies or triopolies, where firms may also set the same price.

on how well they are informed about the price distribution. Moreover, this illustrates that we should not only investigate the effect of entry on average prices, but also on the price distribution, since consumers with different access to information will be heterogeneously affected by entry and profit from changes at different points of the price distribution.

Table 2.2: Descriptives on Prices and Price Dispersion at Market-Date Level

1 km Radius		p10	p25	Mean	p75	p90	S.D.
Mean Price	$\bar{p}_{it}$	103.9	109.4	116.6	124.2	129.2	10.22
Min Price	p_{it}^{min}	102.9	108.9	115.7	122.9	127.9	10.09
Max Price	p_{it}^{max}	104.9	109.9	117.6	125.4	130.9	10.46
S.D.	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(p_{it}-\bar{p}_{mt})^2}$	0.000	0.479	1.095	1.414	2.646	1.186
Range	$p_{mt}^{max}-p_{mt}^{min}$	0.000	1.000	1.856	3.000	4.000	2.031
VOI	$\bar{p}_{mt}-p_{mt}^{min}$	0.000	0.250	0.897	1.250	2.000	0.985
2 km Radius		p10	p25	Mean	p75	p90	S.D.
Mean Price	$\bar{p}_{it}$	104.2	109.6	116.6	124.2	129.1	10.18
Min Price	p_{it}^{min}	102.9	108.9	115.4	122.9	127.9	10.01
Max Price	p_{mt}^{max}	104.9	110.9	118.1	125.9	130.9	10.49
S.D.	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(p_{it}-\bar{p}_{mt})^2}$	0.000	0.577	1.251	1.704	2.748	1.118
Range	$p_{mt}^{max}-p_{mt}^{min}$	0.000	1.000	2.667	4.000	6.000	2.472
VOI	$\bar{p}_{mt}-p_{mt}^{min}$	0.000	0.500	1.239	1.667	2.750	1.167

Note: Dispersion measures are only calculated for market-date combinations with at least two firms. Markets around entrants not included. All values in ct/l. p_{it} denotes a station's i price on date t. p_{mt}^{min}, p_{mt}^{max} and $\bar{p}_{mt}$ denote the minimum, maximum and mean price in market m on date t respectively.

Importantly, the price differences between stations cannot be fully explained by time-invariant differences in station characteristics such as service quality or other measures of vertical differentiation. We show this by documenting how the price ranking between stations changes over time. We implement the rank reversal test proposed by Chandra and Tappata (2011). For any pair of stations A and B, forming a couple c, the test calculates the likelihood that the station which is the cheaper on most days sets the higher price. If station A usually is the cheaper station, then the following formula gives the rank reversal measure:

$$rr_c = \frac{1}{T_c}\sum_{t=1}^{T_c} 1[p_{A,t} > p_{B,t}], \qquad (2.1)$$

where T_c is the number of days on which both stations operate simultaneously, and $p_{A,t}$ and $p_{B,t}$ are the respective prices on day t.

The average rank reversal of pairs with a maximum distance of 1 km (2 km) is 0.105 (0.117). Also, rank reversals increase with the distance between stations. Figure 2.B.3 and Table 2.C.1 in the Appendix show that reversals are lowest for station pairs up to 250 metres distance between the stations, i.e. the likelihood of rank reversal decreases by 5.7 percentage points relative to station pairs between 0.25 and 5km distance. Hence, there is substantial temporal price dispersion at the station-pair level, which cannot be explained by time-invariant differences between stations. Information frictions can explain the observed changes in price rankings.

2.2.2 Market Entry

In Germany, gasoline stations are not subject to any geographic restrictions, and gasoline stations can, in principle, enter almost everywhere. The administrative process is fairly straightforward and common practice, especially for large chains, who are predominant in Germany. As for any business facility, setting up a gasoline station requires a building permit and a business registration in the respective municipality. On top of that, gasoline stations need to satisfy some other requirements, regarding, e.g., ground water safety, fire safety and emissions laws. As such, entry restrictions are viewed as a limited barrier to entry. However, setting up the infrastructure for a new gasoline station is costly and time-consuming, often taking several months.

We define entry as the first time a station posts a price, which should be very accurate given the mandatory price reporting regulation.[11] Based on these entry dates, we allocate treatment to all stations in a 1 km (2 km) radius at the weekly level.

We record over 700 unique entry events in our sample period. Overall, almost 1,300 stations (i.e., approximately 9% of all non-entrants in our sample) experienced at least one entry event within a 2 km radius over our sample period. More than 500 stations experienced entry even within a 1 km radius. In Figure 2.1, we illustrate the geographic and temporal distribution of entry events.[12] Entry occurred all over the country and is fairly evenly distributed over time. Slightly fewer entries in later years might reflect the progressing consolidation of the industry as the number of stations in Germany has been falling over the last decades (Scope Investor Services, 2021).

2.3 Empirical Strategy

We estimate the causal relationship between market entry and station- and market-level outcomes by exploiting the high number of market entry events documented above. In the following, we first explain the estimation setup. Second, we discuss how our identification strategy addresses standard endogeneity concerns when studying market entry. Lastly, we outline the estimation procedure, which allows us to study the heterogeneous effects of competition on consumer types, depending on their information about prices.

We study the effect of the staggered 'rollout' of entries in a two-way fixed effect difference-in-differences model, which compares prices of treated *incumbents* before and after entry with

[11] Stations which changed their brand affiliation are also listed with a new identifier. We do not count these as 'entry events'. For that purpose, we ensure that there is no station of (almost) identical address. Specifically, we check for each 'potential entrant' with a formerly existing station within 400 metres distance whether this station nearby might be the same station under a different affiliation. We check all these cases manually to arrive at a final list of entries. Despite all efforts, we might erroneously classify incumbent stations as entrants. In that case, we are likely to underestimate the price effect of entry in absolute terms.

[12] Figure 2.B.4 in the Appendix provides the distance distribution between entrants and incumbents.

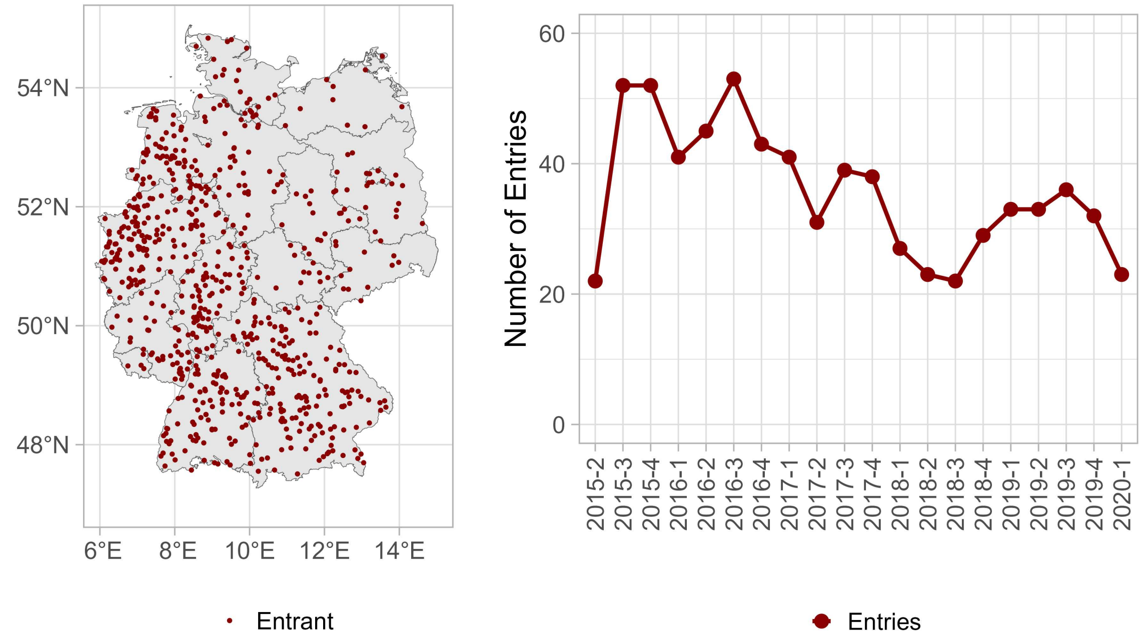


Figure 2.1: Geographical and Temporal Distribution of Entry

Note: The left figure documents the exact locations of all entry events in our sample. The borders represent the administrative federal state borders ($N = 16$ states). The shapefile is provided by ESRI Germany (2025). The right figure plots the distribution of entry events over time aggregated at the quarterly level.

non-treated control stations. The baseline regression is as follows:

$$Y_{it} = \alpha_i + \lambda_{st} + \beta \times 1[Post - Entry]_{it} + \gamma' X_{ct} + \epsilon_{it}. \tag{2.2}$$

Here Y_{it} is station or market i's outcome (e.g., station price or market mean price) on day t. Individual station (or market) and state-times-day fixed effects ($N = 16$ states) are denoted by α_i and λ_{st} respectively. The vector X_{ct}, with c denoting a county, consists of socio-economic variables at the county-year level (e.g., population, unemployment rate, income), which we include as control variables. Most importantly, $1[Post - Entry]_{it}$ is the binary treatment indicator. It will equal one when an incumbent has experienced entry (within 1 km or 2 km radius, respectively) in period t or earlier. The error term is denoted by ϵ_{it}.
We only consider the first entry event for each incumbent and drop observations as soon as a second entry event occur. By doing so, we avoid pooling the likely structurally different entry effects of a first and second entry. Further, we have a clear pre- and post-treatment period for each treated station then. This allows us to estimate the simple difference-in-differences coefficients β, which gives us the effect of entry on outcome Y_{it}, and enables us to implement the staggered difference-in-differences correction by Sun and Abraham (2021) for absorbing treatment on the same sample.[13] We account for correlation in the residuals of the regression by clustering standard errors at the municipality level.

[13]Note that only 2.3% (1 km) and 4.6% (2 km) of all incumbents experience more than one entry event in our sample period. Only looking at the first four months after entry, no station in our sample experiences a second competitor entry.

2.3.1 Identification

Analyzing firm entry is empirically challenging. Entry is part of firms' equilibrium behavior, i.e. entry is endogenously determined by many market circumstances (see, e.g., Berry and Reiss, 2007, Seim, 2006, Toivanen and Waterson, 2005, Schaumans and Verboven, 2015, Bourreau et al., 2021). The econometrician might not observe all market characteristics which determine entry incentives. Comparing *entrants with treated incumbents*, we show that entrants typically have significantly fewer competitors in a 1 km or 2 km radius even when controlling for narrow population data from the RWI-GEO-GRID data (see Table 2.C.2). As profitability and prices are the main driver of entry, a reverse causality bias is a concern. We address this problem by holding such unobserved characteristics of stations and markets fixed over time by including the respective fixed effects and by only exploiting the, from the perspective of the incumbent, exogenous timing of entry.

An additional challenge to identification may be that market-level profitability could change over time. This will be a problem if profitability does not change at the same time and in the same fashion in the treatment and control group, i.e., if treatment and control stations are not on the same trend. To alleviate this concern, we provide event studies for our difference-in-differences analyses estimated based on the following regression form:

$$Y_{it} = \alpha_i + \lambda_{st} + \sum_{\tau=-\underline{\tau},\neq -1}^{\bar{\tau}} \beta_\tau 1[Entry]_{it,\tau} + \epsilon_{it}, \tag{2.3}$$

where $\sum_{\tau=-\underline{\tau},\neq -1}^{\bar{\tau}} 1[Entry]_{it,\tau}$ are leads and lags of the treatment. In our baseline regressions, we bin leads and lags to six-month bins with an effect window of four bins before and five bins after the treatment ($\underline{\tau} = \bar{\tau} = 4$).[14] Endpoints are binned, i.e., the most left and right bins include all observations lying outside the effect window.

If the pre-trends of the event studies are flat, this will be indicative of treated and control stations being on parallel trends. Simultaneously, it will alleviate concerns of a reverse causality bias. In the case of reverse causality, diverging pre-trends should be observable, too. We find that pre-trends are rather flat in all of our event studies later on. Comparing *treated incumbents with all other outside stations before treatment*, we even show that they are very similar in observable characteristics (see Table 2.C.3 in the Appendix). Incumbents do not differ in their price or margin level as well as county-level observables. Markets around incumbents only differ from non-incumbent markets in their higher local population level (3x3 km population data from the RWI-GEO-GRID data). This is suggestive of the estimated effects likely being close to representative of the overall population of stations within the same state-date cell in our sample, too.

Another concern is that we might pick up unrelated shocks, which occur at the same time and in the same markets as the entry events. Then, the estimated treatment effects would be unrelated to the entry events. However, the entry events are geographically dispersed

[14] We later on provide additional short-run analyses with weekly and monthly bins to study incumbents' immediate reactions to entry.

and occur at different points in time. Hence, unrelated shocks would need to take place in a very similar nature in time and space as entry events to invalidate our identification strategy. This seems unlikely. To provide further support that no other demand-side factors change at the exact time of entry, we show that the traffic intensity does not change around the entry event. We use traffic flow data from almost 2,000 traffic counters all over Germany. The left panel of Figure 3.E.4 in the Appendix shows that traffic at counters near entrants does not change around the time of entry.

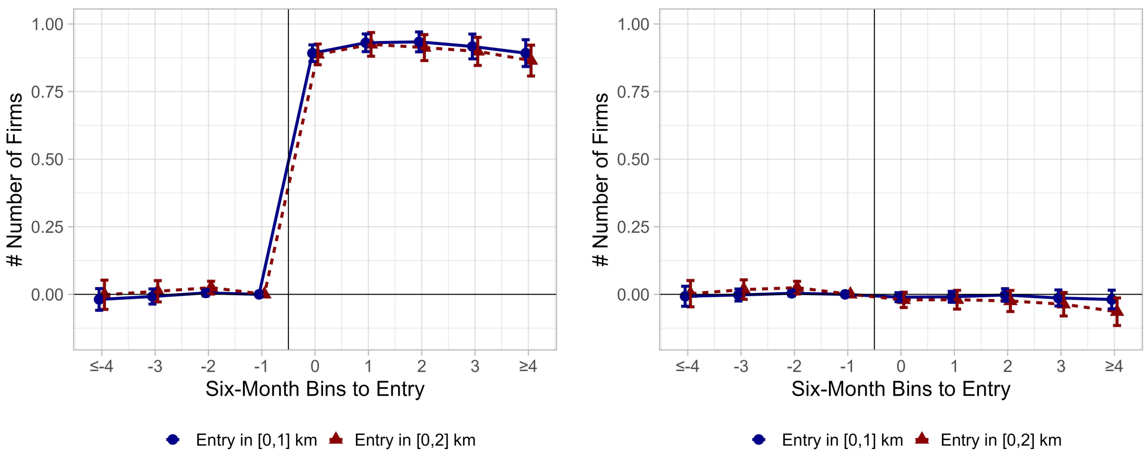
Besides traffic, we study the development of local population numbers as a demand proxy. For this, we make use of annual 1x1 km grid population data from the RWI-GEO-GRID data. For each station, we determine in which grid cell it is located, and calculate the population in the nine grid cells around the station, i.e., a grid of 3x3km. This approach covers an area of 9km^2, which closely matches the area covered by our main estimation approach (3.14km^2 and 12.56km^2 for 1 km and 2 km markets, respectively). The right panel of Figure 3.E.4 shows that the population does not change before or after entry in the area around entrant stations relative to areas around all other stations in Germany. This also holds for a wider 5x5km grid to account for population changes in more distant areas.

Moreover, for our difference-in-differences regressions to uncover the causal relationship it requires that the stable unit treatment variable assumption (SUTVA) is not violated. This implies that spillovers of the treatment of incumbents should not affect control stations. As identification stems from comparing prices within federal states and given the typically very narrow markets in the gasoline industry, it seems unlikely that control stations are strongly affected by entry. If at all, the strategic complementarity in incumbent and control station prices would result in underestimating the effect in absolute terms.

Furthermore, a natural concern is that the studied local shock on competition might not be persistent over time. For example, entry could occur in markets where exit has recently happened or is anticipated. Gasoline markets also have seen a decline in outlets across many countries throughout the last decades (Eckert and West, 2005, Sen and Townley, 2010).[15] However, we provide evidence that the positive shock of entry on the number of firms is persistent. Figure 2.2 shows in its left panel that entry increases the number of stations in the short and long run for more than two years. That the number of firms does not jump up exactly by one is driven by the fact that some stations do not open every day. It is not driven by simultaneous exit, i.e. we do not find that stations are triggered to exit the market in response to entry (right panel). Flat pre-trends also show that entry does not happen in markets that experienced a declining number of firms in advance.

Finally, the recent literature on two-way fixed effects difference-in-differences estimations with a staggered rollout of treatment shows that estimated effects might be biased (Borusyak et al., 2024, Callaway and Sant'Anna, 2021, De Chaisemartin and d'Haultfoeuille, 2020, Sun and Abraham, 2021). This will be the case if treatment effects are heterogeneous for incumbents who experience entry at different points in time. For example, due to changes in the search

[15] The German gasoline market shrunk from 45,000 stations to less than 15,000 stations from 1970 to the early 2000s (Scope Investor Services, 2021). However, since then the number of stations has remained relatively constant.



(a) Including Entrant

(b) Excluding Entrant

Figure 2.2: Market-Level Number of Firms

Note: This figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. The outcome variable is the number of firms in a market on a certain date. In the left plot, entrants are included. In the right plot, only incumbents are considered. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

behavior of consumers over time, effect sizes might vary over time. Our results are unlikely to be qualitatively affected by this potential problem as the vast majority of stations in our sample are never-treated control stations (Borusyak et al., 2024). Nevertheless, we provide robustness checks with the proposed estimator of Sun and Abraham (2021) later on.

2.3.2 Quantile Treatment Effects and Counterfactual Distribution

We want to study how consumers, who hold different information about prices, are heterogeneously affected by station entry. Better-informed consumers purchase at different prices than less-informed consumers. We, therefore, examine how the price distribution changes over its entire range instead of just at the (conditional) mean. As we need to observe pre-entry prices to construct the counterfactual, this analysis focuses on the incumbent prices and does not include entrants' prices.[16]

First, we estimate the quantile treatment effects of entry. Formally, denoting the distribution of incumbents' observed post-entry prices by $F_{entry}(p_{it})$ and the counterfactual distribution of incumbents' prices absent entry by $F_{no\ entry}(p_{it})$, the quantile treatment effect at quantile q is defined by $F^{-1}_{entry}(q) - F^{-1}_{no\ entry}(q)$.

Note that, in our setting, we are especially interested in the *unconditional* quantile treatment effect instead of the *conditional* quantile treatment effect from standard quantile regressions. The latter gives the distributional effects conditional on covariates. For example, the conditional quantile treatment effect at the 90th percentile quantifies the effect on stations with

[16]The latter would not contribute to the identification of the quantile treatment effects. However, they still matter for the consumer's choice and rents in the end, so that we will take them into account in our market-level analysis on market-level minimum, mean and maximum prices later on.

relatively high prices (given their characteristics), but not on stations that are at the 90th percentile of the unconditional price distribution. We, however, care about the general effects along the price distribution - also across stations of different characteristics as consumers likely compare such within a market - and hence are interested in the unconditional quantile treatment effects.

While the formal expression for the quantile treatment effect is rather simple, the econometrician's challenge is to obtain the unobserved counterfactual price distribution $F_{no\ entry}(p_{it})$. We implement two approaches that construct the counterfactual price distribution: first, we estimate quantile treatment effects by using the 'recentered influence function (RIF)' by Firpo et al. (2009). This method uses a local linear approximation of the counterfactual distribution based on observed prices to implicitly derive counterfactual prices. The approximation allows us to estimate quantile treatment effects with OLS regressions. While this provides valuable insights about the change at a chosen quantile of the distribution, it does not allow us to draw conclusions about the global shift of the price distribution. For example, testing whether the distribution of observed prices is first-order stochastically dominated by counterfactual (non-entry) prices is desirable, as this is a sufficient condition for all consumers to profit from entry in a homogeneous goods market. In a second step, we, therefore, elicit the *complete* counterfactual distribution of prices absent entry by adopting the 'distribution regression' approach suggested by Chernozhukov et al. (2013).[17]

Firpo et al. (2009) relies on a local linear approximation of the counterfactual price distribution. The intuition behind the approach is as follows: upon construction of a binary indicator of whether a price observation is above a price threshold p, a simple regression as in (2.2) yields the estimated vertical distance $F_{entry}(\hat{p}) - F_{no\ entry}(\hat{p})$ between the distribution of actual and counterfactual prices at a given price $\hat{p}$. The larger the empirical density at price $\hat{p}$, $f_{entry}(\hat{p})$, the smaller the change in prices (horizontal distance) needed to reach the found vertical distance. This horizontal price distance between the two distributions defines the quantile treatment effect. Formally, this is implemented by transforming the outcome variable to the 'recentered influence function (RIF)':

$$Y_{it}(\hat{q}) = p(\hat{q}) + \frac{\hat{q}}{f(p(\hat{q}))} - \frac{1[p_{it} < p(\hat{q})]}{f(p(\hat{q}))}. \tag{2.4}$$

Here $Y_{it}(\hat{q})$ is the RIF and $p(\hat{q})$ the price at the $\hat{q}^{th}$ quantile. The function $f(p(\hat{q}))$ denotes the empirical density function of observed prices evaluated at $\hat{q}$. We estimate $f(p)$ using an Epanechnikov kernel with a bandwidth following Silverman's rule of thumb as in, for example, Dube (2019). Given $\hat{q}$, $p(\hat{q}) + \frac{\hat{q}}{f(p(\hat{q}))}$ is constant, so the main variation in the new outcome variable stems from whether the price is below or above $p(\hat{q})$. Estimating (2.3) with the RIF as the dependent variable allows us to obtain the quantile treatment effect at a certain quantile in only one regression. Applying this approach for different quantiles $\hat{q}$ allows us to

[17] The estimation of distributional effects of public policies is especially common in the field of public and labour economics. For example, see Havnes and Mogstad (2015) and Huebener et al. (2017) for applications of Firpo et al. (2009). Dube (2019) and Hernæs (2020), among others, use the method of Chernozhukov et al. (2013).

estimate the shift at several discrete points of the entire price distribution.
Note that, as we insert (2.4) as the dependent variable in the regression equation (2.2), we identify changes in the empirical distribution within stations/markets over time.[18]
Chernozhukov et al. (2013) propose to construct the counterfactual price distribution by estimating how entry affects the likelihood of prices being below a certain threshold. In our context, the respective regression equation is:

$$1[p_{it} \leq \hat{p}] = \alpha_i + \lambda_{st} + \beta \times 1[Post - Entry]_{it} + \gamma' X_{ct} + \epsilon_{it}. \tag{2.5}$$

The outcome variable is a dummy which will be one if $p_{it} \leq \hat{p}$. The coefficient β corresponds to the estimated vertical difference between the two distribution functions, $F_{entry}(\hat{p})$ and $F_{no\ entry}(\hat{p})$ at price $\hat{p}$. Estimating one regression for each unique price value in the dataset constructs the counterfactual empirical price distribution $\hat{F}_{no\ entry}(p_{it})$. Hence, this approach is computationally more intense but explicitly models the counterfactual distribution instead of a local linear approximation.
In our empirical analysis in Section 2.4, we implement both methods laid out here. We build on the approach by Firpo et al. (2009) to estimate quantile treatment effects. Additionally, we estimate the price counterfactual distribution using Chernozhukov et al. (2013) to then test for a global shift of the price distribution in a first-order stochastic manner, as elaborated in the following subsection. The outcome from Chernozhukov et al. (2013) also serves as an input to cross-validate our results when using the method of Firpo et al. (2009) instead, by comparing the empirical price distribution and the counterfactual price distribution.

2.3.3 First-order Stochastic Dominance

The quantile treatment effects are, by design, informative about the effect of entry at certain points of the price distribution *in isolation*. However, we are also interested in how entry shifts the entire price distribution. For example, we can ascertain that consumers who are very well-informed about prices benefit from entry only if the price distribution is first-order stochastically dominated by the counterfactual (no entry) price distribution. This requires to *jointly* test for the difference between the distribution of observed and counterfactual prices at multiple prices.
We, therefore, implement three statistical tests to examine whether observed prices are jointly first-order stochastically dominated by counterfactual prices. The one-sided Kolmogorov-Smirnov test is based on the maximum vertical distance between two distributions. In contrast, the stochastic-dominance tests proposed by Barrett and Donald (2003) and Davidson and Duclos (2000) are based on the minimum vertical distance between two distributions and therefore more conservative. To compute those distances, the test by Davidson and Duclos (2000) takes a pre-specified set of points on the price grid as an input. Barrett and Donald (2003) integrate over the entire support of prices, which eliminates the necessity to fix a set

[18] While many markets only have two or three stations, we can still obtain the quantile treatment effects at all quantiles as identification stems from the price comparison of treated stations on any pre- to any post-entry date. State-date specific demand or costs shocks are controlled for by the fixed effects.

of points upfront.

To examine the robustness of our findings, we present the results for all three tests. We compare observed post-entry prices with the counterfactual no-entry prices, elicited through the distribution regressions as in Chernozhukov et al. (2013) as explained above.[19] Note that the elicited counterfactual distribution, given by the point estimates of the distribution regressions, may not be monotonically increasing at all prices. The distribution regressions at all prices do not require a positive density function across distribution regressions. However, this does not affect the test results, as the tests compare the observed price distribution and the counterfactual price distribution at certain prices only. As we take the point estimates of the estimated counterfactual distribution, we abstract from its confidence interval.

2.3.4 Effects on the Value of Information (VOI)

The methods above allow us to estimate the changes along the price distribution. Importantly, this reflects only *posted prices*, which are different from the prices consumers actually pay (*expected transaction prices*). A useful concept to capture this disparity is the *Value of Information* (VOI), defined as difference between the expected price and the expected *minimum* price in a market. VOI can then be interpreted as the extent to which fully informed consumers are better off relative to uninformed consumers, purchasing from a station at random. An alternative interpretation is how much a yet-uninformed consumer would gain if she were to exercise effort to become fully informed (see Varian, 1980, Armstrong et al., 2009, Lach and Moraga-González, 2017, and Honka et al., 2019, for a recent survey through the vast literature on consumer search).

To investigate this effect, we present event studies similar to our baseline specification above, using either VOI or the range of prices as the dependent variable.

2.4 Results

In this section, we first show that, indeed, prices *on average* decrease in response to entry. Building on this result, we then study the heterogeneous effects of entry along the entire price distribution and implications for consumers through the Value of Information (VOI).

2.4.1 Average Effect on Prices

In the first step, we estimate the dynamic effect of entry on posted prices at stations nearby as outlined in equation (2.3). We estimate the event studies for entry happening in a radius of 1 km or 2 km around a station. The left panel of Figure 2.3 shows that we find a negative effect of entry on prices. Flat pre-trends suggest that we do not violate the parallel trends assumption and, by that, are indicative of a causal relation between entry and prices. They are also suggestive of no anticipatory pricing in the months before entry because incumbents'

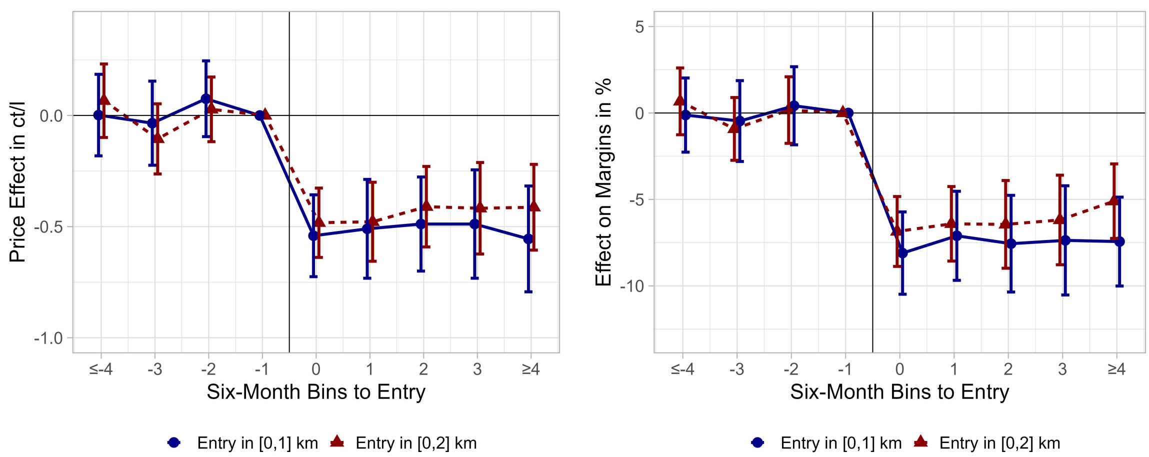
[19] An alternative would be to compare observed pre-entry and post-entry distributions. Since several additional determinants of prices, e.g., input prices, might have changed meanwhile as well, this would be a potentially misleading comparison though.

prices right before entry do not significantly differ from prices more than two years before entry. The price effects are rather persistent over time for entry up to 2 km distance.

When entry occurs close by (i.e., up to 1 km distance between entrant and incumbent), the estimated average effect is a price reduction of around 0.5 ct/l. To put these numbers into perspective, recall that the major brand premium is around 1 ct/l, so only around twice as big as the effect of a single entry. Moreover, a retail price reduction of 0.5 ct/l implies a reduction of around 7% of gasoline stations' gross margins. In the right panel of Figure 2.3, we provide event-study results on the effect on margins explicitly confirming this. In the cross-sectional study by Lach and Moraga-González (2017), a 10% increase in the number of stations is associated with a 0.06 ct/l price reduction in 1 km-radius markets. Since a single entry event increases the number of stations considerably in relative terms (e.g., by 50% and 20% for markets with two and five incumbent stations, respectively), our implied effects are much stronger than those in Lach and Moraga-González (2017).

As expected, the effect of entry is weaker for entry farther away from the incumbent. When pooling entry events in up to 2 km distance around incumbents, the price effect estimates are slightly smaller in absolute terms.

The respective pooled difference-in-differences estimates coinciding with the event studies can be found in Table 2.C.4 of the Appendix. In the Appendix, we also provide more detailed results on how the entry effect declines with increasing distance between entrant and incumbent (see Figure 2.B.6).



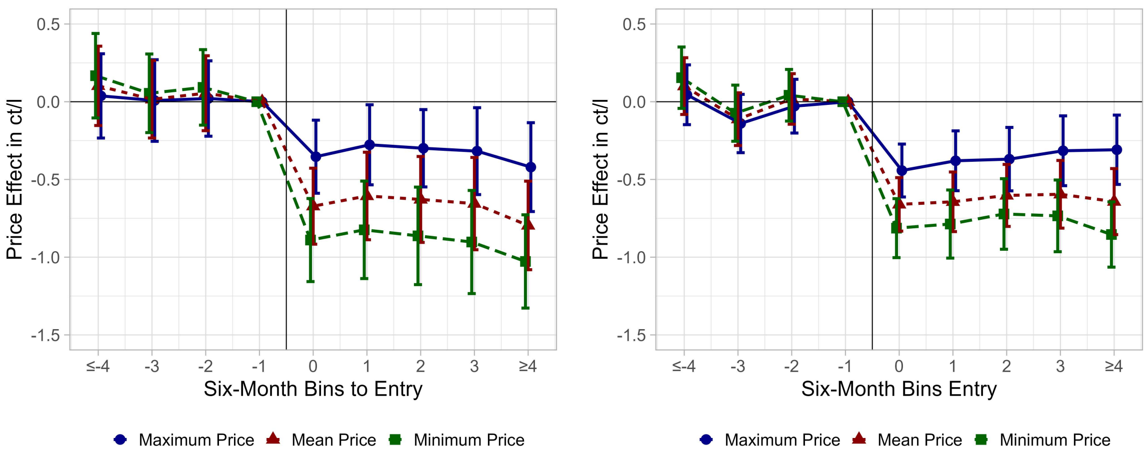
(a) Prices (b) ln(Margins)

Figure 2.3: (Average) Effect of Entry on Prices and ln(Margins)

Note: This figure presents the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

2.4.2 Distributional Effects on Prices

We continue by analyzing the effect along the market-level price distribution. For this, we first show how market-level minimum, mean and maximum prices are affected by the entry



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.4: Effect of Entry on the Price Distribution

Note: This figure presents the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Outcome variables are the market-level minimum, mean and maximum price for each market-date observation. Only observations with at least two firms in the market included. Entrants' prices are included in the calculation. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

events. Subsequently, we investigate the effect of prices at all quantiles of the price distribution and estimate quantile treatment effects.

Figure 2.4 demonstrates that the minimum, mean, and maximum market prices are impacted differentially by entry.[20] For example, the maximum price decreases by less than 0.5 ct/l in response to entry up to a distance of 1 km. In contrast, the mean price drops by 0.7 ct/l and the minimum price falls by 1.0 ct/l. The same qualitative pattern holds for markets of 2 km radius.

We draw the following conclusions from these findings: First, a decrease in the minimum price is indicative of fully informed consumers likely profiting more from entry. Similarly, a decreasing mean price suggests that uninformed consumers, who buy from a random station in the market, also benefit, albeit to a lesser extent than fully informed consumers. Second, the effect of entry on consumer prices is persistent over time.

To get an impression of how the price distribution changed at all prices instead of just the minimum, mean or maximum price, we estimate quantile treatment effects as in Firpo et al. (2009). We do this at every 5th percentile between the 10th and 90th percentile. Figure 2.5 presents the estimation results. While the average price effect of entry in a 1 km radius around a station was approximately 0.5 ct/l, there is substantial variation in the price effect along the price distribution. Quantile treatment effects range from 1.0 ct/l at the left tail of the price distribution to 0.0 ct/l at the right tail of the distribution. For higher quantiles between the 65th and 85th percentile, the effect is less pronounced. A similar pattern is

[20]Only market-date observations with at least two stations are included, since the minimum, mean, and maximum price are identical in markets with just one station.

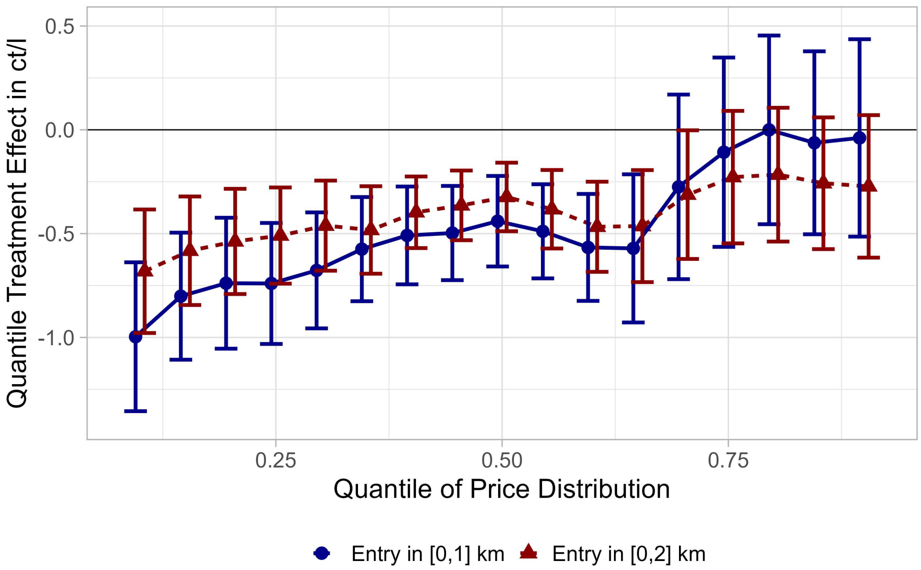


Figure 2.5: (Unconditional) Quantile Treatment Effects of Entry on Prices

Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). Estimates for every 5th percentile between the 10th and 90th percentile are provided. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls are included. Entrant observations are not included. Results obtained from regressions in which we also control for daily, region-specific wholesale price variations deliver almost identical estimates (see Figure 2.B.7 in the Appendix).

observable for entry in a radius of 2 km, although the distributional heterogeneity is less pronounced.

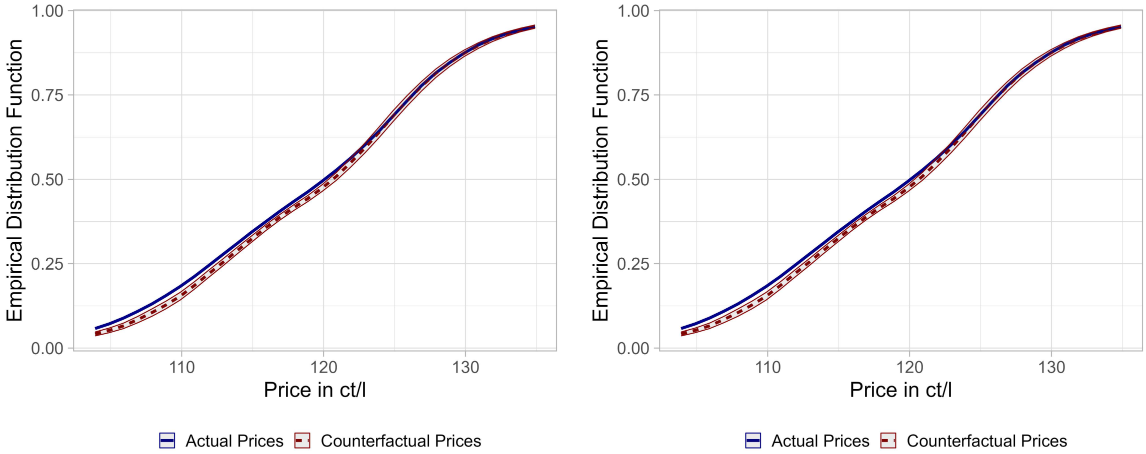
Hence, prices at the left tail of the distribution respond more strongly to entry. This favors more informed consumers as they are more likely to fuel at a low price. Nevertheless, uninformed consumers also profit from low prices, as the expected price decreases.

We cross-validate our results on heterogeneous price effects along the price distribution by eliciting the counterfactual price distribution as in Chernozhukov et al. (2013). With the full counterfactual distribution at hand, we can inspect the difference between observed and counterfactual post-entry prices over the entire support. Figure 2.6 supports our findings that entry affects incumbents' prices especially at the left tail of the distribution. Counterfactual prices (i.e., incumbents' prices after the timing of entry if entry had not happened) are significantly higher than observed prices at the left tail of the distribution.

2.4.3 First-order Stochastic Dominance

As Figure 2.6 shows, entry shifts the whole price distribution upwards. If the price distribution is shifted upwards at each percentile, this would be indicative of all consumers, independent of their information about prices, being positively affected by entry. We, thus test whether observed post-entry prices are first-order stochastically dominated by counterfactual prices.

We provide results for three different tests: the one-sided Kolmogorov-Smirnov (KS) Test,



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.6: Observed Prices and Counterfactual Distribution

Note: The figures plots the actual price distribution of incumbents' prices, which experienced entry (blue) and the estimated counterfactual price distribution (red) of a scenario without (or before) entry. The counterfactual distribution is obtained using distribution regressions as proposed by Chernozhukov et al. (2013), i.e. the regressions as in equation (2.5) for price thresholds at each integer ct/l. The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

as well as the two tests proposed by Davidson and Duclos (2000) and Barrett and Donald (2003). The one-sided KS-Test builds on the maximum, one-sided distance between the distribution of observed and counterfactual prices. The null hypothesis of equal distributions is tested against the alternative that the counterfactual distribution is smaller than the observed prices. However, for this test only the maximum distance and not the distance between the distributions over the full support matters. The tests by Davidson and Duclos (2000) and Barrett and Donald (2003) examine whether the counterfactual distribution dominates the empirical distribution function of observed prices at multiple prices at the same time. In both tests, the null is first-order stochastic dominance. The alternative in Davidson and Duclos (2000) is no restriction and in Barrett and Donald (2003) that observed prices first-order stochastically dominate the counterfactual prices.

Test results can be found in Table 2.3. The one-sided KS-Test rejects equal distributions for the one-sided alternative. Both other tests do not reject that entry indeed shifts the price distribution in a first-order stochastic manner upwards. The Davidson and Duclos (2000) test does not reject first-order stochastic dominance if all test statistics calculated at 17 equidistant prices between the 5th and 95th quantile of the observed price distribution lie below the critical value of 3.963 (significance level 1%) and if at least one absolute value of a test statistic exceeds the critical value. Indeed, this is the case for our 1 km and 2 km markets. For the test of Barrett and Donald (2003), we report p-values of rejecting the null hypothesis based on comparisons of the two empricical CDFs (ECDFs) at every integer ct/l between the 5th and 95th quantile of the observed price distribution. For both 1 km markets and 2 km markets, the null hypothesis of first-order stochastic dominance clearly cannot be rejected. For 1 km markets, the p-value is 0.93. For 2 km markets, the p-value is 1, since the counterfactual CDF is smaller than the observed CDF at literally all examined prices.

Table 2.3: First-Order Stochastic Dominance

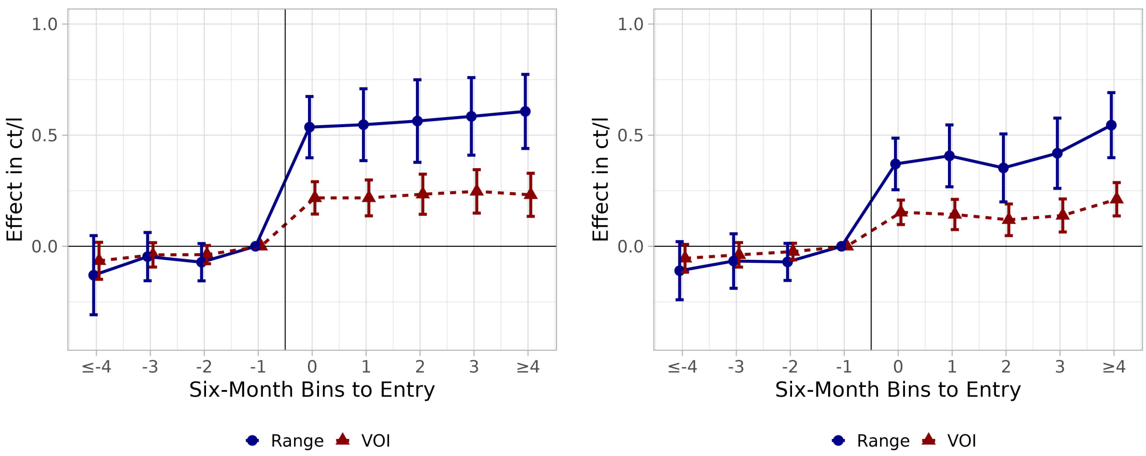
Entry in [0,1] km	H_0	Test-Statistic	p-value (H_0 reject)
One-Sided KS-Test	$F_{no\ entry} = F_{entry}$	0.0292	< 0.01
Davidson and Duclos (2000)	$F_{no\ entry} \leq F_{entry}$	$\in [-40.575, 0.915]$	-
Barrett and Donald (2003)	$F_{no\ entry} \leq F_{entry}$	0.195	0.93
Entry in [0,2] km	**H_0**	**Test Statistic**	**p-value (H_0 reject)**
One-Sided KS-Test	$F_{no\ entry} = F_{entry}$	0.0201	< 0.01
Davidson and Duclos (2000)	$F_{no\ entry} \leq F_{entry}$	$\in [-46.085, -8.633]$	-
Barrett and Donald (2003)	$F_{no\ entry} \leq F_{entry}$	-	1

Note: This table reports the results of tests of first-order stochastic dominance. The tests differ in their design. The one-sided KS-Test tests the null hypothesis against the alternative that $F_{entry} > F_{no\ entry}$. It evaluates the maximum distance between the actual and counterfactual distribution. Davidson and Duclos (2000) extend the comparison of both distributions to more than the point of maximum distance. We implement the FOSD test by Davidson and Duclos (2000) as done in Asplund and Nocke (2006) and Gavazza (2011). Barrett and Donald (2003) smooth out the comparison by integrating over several points. Hence, the comparison can be interpreted as more continuous along the distributions. We evaluate the Davidson and Duclos (2000) approach at all prices in equal distance between 1.029 and 1.349 Euro/l, which approximately represents the 5th and 95th percentile of the distributions in steps of 2 ct/l (17 points). For Barrett and Donald (2003), we take the supremum of the distance between both ECDFs at every full ct/l between 1.029 and 1.349 Euro/l. The critical value at a significance level of 0.01 for the Davidson and Duclos (2000) test comes from the studentized maximum modulus distribution (Stoline and Ury, 1979). Given 17 equidistant points at which we calculate the test statistic, the critical value at the 1% significance value is 3.963. All test statistics have to be below this value and at least one absolute value of a test statistic needs to exceed the critical value. No test statistic is reported for Barrett and Donald (2003), as the ECDF of counterfactual prices is smaller than the ECDF of observed prices at all examined prices, and thus the null hypothesis clearly cannot be rejected.

2.4.4 Effects on the Value of Information (VOI)

In this subsection, we show how the Value of Information (VOI) is affected by entry. This captures the saving potential for consumers through expected *transaction prices* instead of *posted prices*. Consumers with relatively easy access to information, or those who are very eager to engage in price comparison, will purchase at the expected *minimum*. The relative gain of doing so instead of purchasing at average prices is measured by VOI.

The effect of entry on the VOI is shown in Figure 2.7. The value of information (VOI) increases by about 0.25 ct/l after entry in 1 km radius markets. Relative to the pre-entry average VOI, this corresponds to a 29% increase. The effect is persistent over time throughout our sample. The effect on the *range* of prices is even stronger, as it increases by 0.6-0.7 ct/l (i.e., by around 35%) after entry in 1 km radius markets. Consistently across measures and market delineations, we observe an increase in price dispersion. Post-entry, expected prices



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.7: Market-Level Dispersion Effects of Entry

Note: This figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Only observations with at least two firms in the market included. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

decrease, but expected *minimum* prices decrease even stronger, leading to an increase in VOI and the range.

Using quantile treatment effect methods as above, we further show that the increase in dispersion is evident along the entire dispersion distribution (see Figures 2.B.8 and 2.B.9 in the Appendix). Hence, the value of information increases in all markets, although the effect is especially pronounced in formerly less dispersed markets. These findings are different from Moraga-González et al. (2017) who did not find increasing competition to be associated with higher dispersion. In contrast, Chandra and Tappata (2011) find more rival firms to increase market-level dispersion.

Importantly, our measure through VOI is likely to be a lower bound on the gains of entry for consumers. The reasons are the following. Throughout our main analysis, we exclude the entrants' prices and consider incumbent prices only. However, consumers certainly do (weakly) gain additionally from the presence of an additional firm. Entry additionally effects consumers' propensity to search for lower prices. An additional supplier brings benefits for consumers, as it may provide easier access to more prices. This means that the effective number of prices observed by each consumer, for example during their daily commute, increases due to higher station density and reduced search costs (Dubois and Perrone, 2017). The search incentives of consumers also change as an additional firm in the market affects the resulting price dispersion in the market. Given that we showed that price dispersion, and consequently, the incentive to search, increases with entry, both of these factors suggest that consumers are more inclined to compare prices. Absent a fully specified model about how the distribution of consumer information materializes, we cannot quantify this effect on transaction prices for different consumer types. We therefore take the rather conservative position that the number of price quotes observed by each consumer remains constant. In that sense, we are likely to underestimate the benefits of entry to consumers. Similarly, we abstract from aggregate demand effects due to lower prices. Although the demand elasticity in gasoline markets (Bento et al., 2009, Coglianese et al., 2017, Davis and Kilian, 2011, Li et al., 2014, Kilian and Zhou, 2024) or other real-time pricing markets (Fabra et al., 2021) is very low, this also contributes to us underestimating the overall welfare gains of consumers.

2.4.5 Exit

All our previous results so far examine the effects of station entry. It is instructive to see whether the effects are similar for exit. We therefore repeat the analysis for the mirroring scenario of station exit.

We find that an exit event increases incumbents' prices by 0.332 (0.336) ct/l in 1 km (2 km) radius markets (see Figure 2.D.2 in the Online Appendix for the dynamic effect). In line with our results on entry, we find that a change in competition affects prices the strongest in the left tail of the price distribution (see right panel of Figure 2.D.2). While the results for entry and exit are qualitatively identical, the effects of exit are slightly smaller in absolute terms. This stems from the fact that markets in which entry and exit take place are structurally different from each other. For example, there are on average 2.1 stations in a 2 km radius around entrant stations, but 3.8 stations around exiting stations. This suggests that additional effects, beyond the scope of this paper, arise under exit.

2.4.6 Alternative Mechanisms & Heterogeneous Effects

In this subsection we first discuss and examine alternative channels via which stations could react to gasoline entry other than the pure price mechanism modeled above. We then examine whether the effect of entry is heterogeneous across different types of incumbents and entrants. The corresponding figures can be found in the Online Appendix.

Other Strategic Responses. We investigate whether stations react to entry by adapting non-price instruments such as opening hours in treated markets after entry. If opening hours, as proxy of service quality and automation at the station, are not affected, this indicates that the firm's main responses to entry are price adjustments due to its low menu costs. Figure 2.D.3 shows how entry changes opening hours. To measure opening hours, we construct a daily station-level dummy that is one if a station opens 24/7. Incumbents' opening hours are not significantly different after entry in comparison to before entry. We take this as evidence that firms do not react by adjusting quality along this dimension (Armstrong and Chen, 2009).

Finally, we examine another channel via which entry could have affected prices beyond the pure price pressure of consumers facing another alternative. In Section A, we test how consumer information frictions change in response to entry by extending the test from Chandra and Tappata (2011) to a dynamic setting. Decreasing information frictions might also reduce incumbents' prices. We calculate quarterly, station pair rank reversal measures for pairs of incumbent stations. In an event study setting similar to equation (2.3), we examine how rank reversals are affected when entry occurs near one of the stations of the couple. We do not find changes in information frictions (see Appendix A), which also suggests that the information distribution of consumer types does not change significantly with entry.

Price Cycles. Price cycles, reminiscent of the Edgeworth cycles (Maskin and Tirole, 1988), are a well-documented feature in gasoline markets in many countries, such as the US (Lewis, 2008), Canada (Noel, 2007, 2009, 2019), Australia (Byrne and De Roos, 2019), Norway (Foros and Steen, 2013), and also in Germany (Siekmann, 2017, Assad et al., 2024). Note that in our main analysis, we deliberately focus on daily prices at 5 pm to eliminate inter-temporal price effects generated through price cycles. To ensure that our results do not merely reflect an effect of entry on the price cycle, we first consider the frequency of price adjustment, measured by the absolute number of daily price changes per gasoline station, and the median price change, a measure of cycling asymmetry. The results are shown in Figure 2.D.4. In both of these measures, there is no entry effect, suggesting that entry did not alter the cycling behavior of stations in any significant way.[21] Furthermore, in Figure 2.D.5, we demonstrate that the average effect of entry on prices and dispersion by hour does not vary significantly throughout the day. This evidence suggests that entry does not affect the nature of price cycles.

Algorithmic Pricing. According to Assad et al. (2024), many German gasoline stations have adopted algorithmic pricing software since early 2017, especially larger chains. They show that station prices are affected, especially when multiple stations in a market adopt the technology. To ensure that our results are not driven by the coincidence of entry and technology adoption, we separately analyze the period before early 2017, when almost no station had implemented pricing software already. Figure 2.D.6 shows that both the average treatment effect as well as the distributional implications remain the same.

Heterogeneous Treatment Effects. We examine heterogeneity in the average and distributional effects of entry. We exploit the richness in the entry events across incumbent and entrant characteristics, e.g. brand affiliation or exposure to local competition. Figure 2.D.7 shows heterogeneous effects on prices as estimated in equation (2.2). Note that, as before, we focus on incumbent prices throughout. We first consider heterogeneity regarding observable characteristics of the *entrant*. Prices decrease more strongly when low-price stations such as JET enter. We next investigate heterogeneity regarding observable characteristics of the *incumbent*. We find stronger price effects for brands that do not belong to the Big Four (Aral, Shell, Total, Esso), and similarly, for formerly cheap stations. Both the entrant-specific and the incumbent-specific results are consistent with branded stations offering some additional amenity such as faster service (Png and Reitman, 1994), or are simply perceived as offering

[21] Other studies like Noel (2007) and Siekmann (2017) show that the existence of Edgeworth cycles or their characteristics depend on the local competitive environment. However, these studies compare cycles across markets or stations and do not exploit within-station changes in the competitive environment.

higher quality (Png and Reitman, 1995, Martin, 2024). These branded stations face less elastic demand and respond less aggressively to entry. Finally, there is a stronger price effect in markets in which the entrant brand did not operate a station before.[22]

We further analyse whether the price effect of entry differs by the degree of automation of incumbent and entrant. Automation was linked to prices in gasoline markets outside of Germany (Soetevent and Bružikas, 2018, Kim, 2018). Note that our analysis on automation may be less relevant in Germany than in several other countries since Germany has a relatively low share of unmanned stations. As we do not observe the degree of automation of a certain station directly, we classify stations into automated and non-automated based on their brand affiliation and names. Some chains only operate automated, unmanned stations and automated stations are named as such explicitly. We do not identify clear differences in the price reaction of automated and non-automated gasoline stations. Entry of automated and non-automated stations does not cause different price reactions either. Hence, no different implications for entry interacted with automation arise.

We show that the named dimensions of heterogeneity do not change the qualitative conclusions regarding the distributional effects, but the magnitude of these effects. Exemplarily, we look at entry by the low-price brand JET.[23] We find that low-price entry substantially decreases prices in the left tail of the price distribution (see Figure 2.D.8). At higher prices, quantile treatment effects are not statistically different from zero. This is also supported by the distributional analysis as in Chernozhukov et al. (2013) (see Figure 2.D.9). Correspondingly, the distributional implications of prices also hold for non-JET entries, but in a less pronounced way (see Figures 2.D.10 and 2.D.11). We, further, show in Figure 2.D.12 that the market concentration determines how strong price reactions are at different quantiles. However, the pattern of stronger price effects at the left tail is robust across markets with a different number of non-entrant stations.

2.4.7 Robustness Checks

We continue in describing a set of sensitivity checks to prove the robustness of our estimation results.

Alternative Distributional Analysis. We study the heterogeneous effects of entry along the price distribution by applying an alternative approach to identify heterogeneous effects of a treatment along a distribution of an outcome variable - the method proposed in Cengiz et al. (2019). In separate regressions, we estimate how the treatment affects the likelihood of observing prices in certain price bins of the distribution. Figure 2.D.13 supports our results from above. Entry increases the probability of observing prices in the lowest price bins while reducing the probability of observing very high prices. Observing intermediate price bins is not significantly more likely than absent treatment.

[22]Entrants are of the same brand as incumbents in 2-2.5% of all entry events (depending on whether markets are defined in 1 km or 2 km radii).

[23]Prices by brand are reported in Figure 2.D.1.

Identification Cells & Inference. We test whether our results are robust to different identification cells, i.e. identification within different layers of (non-)administrative regions. Figure 2.D.14 presents the baseline event studies for including Date, County-Date (401 counties) or ROR-Date (96 'Raumordnungsregionen') fixed effects instead of State-Date (16 states) fixed effects[24]. A narrow identification grid reduces the average price effect of entry. Though, the concern of spillovers within narrow geographical areas of counties might downward-bias the effect size (in absolute terms). The qualitative results hold for both market definitions. Also, the quantile treatment effects are similar for all fixed effects combinations as evident in Figure 2.D.15.
We, moreover, check whether the interpretation of our results changes when standard errors account for correlation in residuals within stations, within counties or RORs instead of within municipalities. Our results are unaffected by this (see Figures 2.D.16 and 2.D.17).

Binning. We show that the average estimated price effects are not sensitive to the choice of the number of leads and lags. Figure 2.D.18 shows that the event studies look very alike for different approaches. We further report an event study (see Figure 2.D.19) which zooms in the weeks and months right around the entry events. This figure also illustrates that the main effect of entry materializes rather quickly, i.e., within a two-weeks time window.

Market Size. To ensure that our effects are not just driven by very small, concentrated markets, in which entry might have the largest impact, we also estimate the distributional analysis for markets with at least four stations. In doing so, we focus on 2 km market only as the majority of 1 km markets have less than four active stations. Our results remain qualitatively the same (see Figure 2.D.20).
Further, we restrict our baseline market-level analysis to markets with at least two stations. If entry into monopoly market triggers exit, we would omit such markets as only one station remains. We, therefore, rerun our market-level analysis based on a sample where we include treated markets based on their pre-treatment concentration instead of by date-specific concentration. Figure 2.D.21 shows that our results remain unaffected if we only include treated markets with at least two stations in the market in the last pre-entry quarter.

Staggered Difference-in-Differences. Recent work by Borusyak et al. (2024), Callaway and Sant'Anna (2021), De Chaisemartin and d'Haultfoeuille (2020) and Sun and Abraham (2021) has shown that treatment effect heterogeneity in combination with staggered treatment can bias the estimated treatment effect. In our setting, changes in search behavior over time might change the dynamic treatment effect across stations that are treated at different points in time. While the large number of never-treated units mitigates this concern (Borusyak et al., 2024), we further address this concern by exemplarily also running our main regressions with the estimator proposed in Sun and Abraham (2021). The estimates support our qualitative findings. Figure 2.D.22 shows that entry (on average) decreases prices in a similar fashion as in Figure 2.3. If at all, the pattern becomes more extreme with no price

[24] We match stations to the different levels of regions using shapefiles by Bundesamt für Kartographie und Geodäsie (2025*a*,*b*).

effect at all for the top five deciles. Moreover, the right-hand side of the figure clearly shows that quantile treatment effects are stronger at the left tail of the price distribution. This also is in line with the differences between observed and counterfactual prices as reported in Figure 2.D.23.
Beyond that, treatment effects of entry may vary with the number of entries which take place near an incumbent. In our analyses, we decided to only focus on the first entry which an incumbent experiences (s. Section 2.3). In Figure 2.D.24, we now further show that our results look identical when allowing for multiple treatments per incumbent, so that the effects of consecutive entry events are treated as additive effects.

Alternative Daytimes & Types of Fuel. A concern in our main analysis might be that our results are sensitive to the time of the day or the type of fuel, e.g. when the composition of consumers changes over the course of the day or demand elasticity varies between fuel types (Montag, Sagimuldina and Winter, 2023). Hence, we carry out the analysis also for 7 am and 12 am as well as for E5 gasoline prices instead of diesel. Figures 2.D.25 and 2.D.26 show, on the one hand, that results on diesel fuel are not strongly affected by the time of the day. On the other hand, there is no qualitative difference between our results for diesel and E5 gasoline.

2.5 Conclusion

In this paper, we study station entry in the German gasoline market and its impact on prices and the implications for different consumer types. Owing to our comprehensive data set with a very high number of unique entry events, we thus complement existing findings in the literature regarding the *average* effect of prices. The average effect of entry is a price reduction of around 0.5 ct/l. For the firms operating in that market, this reduction is certainly non-negligible as it corresponds to a reduction in gross margins by 7%. Since entry also affects the gains of search and thus the distribution of consumer information types, this is likely to be a conservative estimate of the effect of entry.
Additionally, we find that entry decreases prices in a first-order stochastic dominance manner. However, price effects are stronger at the left tail of the distribution. These changes along the price distribution have heterogeneous effects on consumers, depending on how well informed they are about prices. We show that entry increases the Value of Information (VOI) by 29%, implying that informed consumers benefit considerably more from entry. Uninformed consumers, who already pay higher prices to start with, do benefit from entry, but less so than other consumer types. Thus, our results have important implications for the distributional effects of competition on different types of consumers in markets with information frictions.
In this paper, we do not take a position regarding which segments of the population tend to fall into the informed or uninformed consumer category, respectively (Byrne and Martin, 2021). Irrespective of whether high-income consumers have more or less elastic demand, they may be more inclined to search (Nishida and Remer, 2018, Fischer et al., 2024), rendering them the main beneficiaries of entry. In that case, policies regarding entry barriers and

market structure also need to take into account distributional consequences. Through this channel, our findings inform policymakers to better understand the relevance of competition for different consumer groups.

Appendix A: Information Frictions

Price dispersion in homogeneous goods markets can arise from the fact that consumers hold heterogeneous information (Varian, 1980). However, entry could improve consumer information about prices. More firms in a market imply a higher likelihood to observe prices (e.g., for commuters). Hence, entry may decrease information frictions. Chandra and Tappata (2011) proposed a static test for information frictions based on how often the price ranking of a pair of neighboring stations alters. The intuition is that rank reversals in a homogeneous goods market can hardly be rationalized without information frictions. For a detailed explanation of the static test see Section 2.2.

To examine how price rank reversals change over time and after entry, we extend their environment to a dynamic setting. Specifically, we calculate a rank reversal measure for each station pair and quarter. Hence, the outcome variable - the measure of rank reversal between station A and B (couple c) in quarter q where station A sets the lower price on at least half of all days in a quarter - is constructed in the following way:

$$rr_{c,q} = \frac{1}{T_{c,q}} \sum_{t=1}^{T_{c,q}} 1[p_{A,t} > p_{B,t}], \tag{2.6}$$

where $T_{c,q}$ is the number of days in quarter q on which both stations report a price.

We then take $rr_{c,q}$ as the outcome in the following event-study regression:

$$rr_{c,q} = \alpha_c + \lambda_{sq} + \sum_{\tau=-\underline{\tau},\neq -1}^{\bar{\tau}} \beta_\tau 1[Entry]_{cq,\tau} + \epsilon_{cq} \tag{2.7}$$

where we include couple fixed effects (α_c) and state-quarter fixed effects (λ_{sq}).[25] $1[Entry]_{cq,\tau}$ gives the leads and lags of entry events. A couple will be treated if at least one station experiences entry in a given radius. Similar to our approach in the main analysis, we only look at the first entry event per couple and drop observations after the date of the first entry. Ex-ante the effect of entry on rank reversals is ambiguous. Entry might foster rank reversals because, e.g., a new entrant very close to an incumbent might make the price ranking with other stations less important. At the same time, entry might improve consumers knowledge about prices. Then, information frictions and rank reversals could decrease. Figure 2.A.1 provides results of entry effects on rank reversals. We report the effect of entry in a 0.5 km, 1 km and 2 km radius.[26] In all three cases, we cannot identify a change in the likelihood of rank reversal. Our results are also barely sensitive to whether looking at entry's effect on couples with a bilateral distance of up to or above 1 km (see Figure 2.A.2) or only looking at entry events where the entrant is the nearest competitor of at least one of the stations (see Figure 2.A.3). We take this as suggestive for entry not sufficiently affecting information

[25] Note that for some couples, stations might not be located in the same federal state. Hence, we construct the point in the middle between both stations and match this point to federal states. A similar procedure allows us to match county-level covariates to the regression and allows for clustering at the municipality level.

[26] We add the more narrow entry radius of 0.5 km as Chandra and Tappata (2011) suggest that rank reversal of very close stations is significantly lower than for other station pairs.

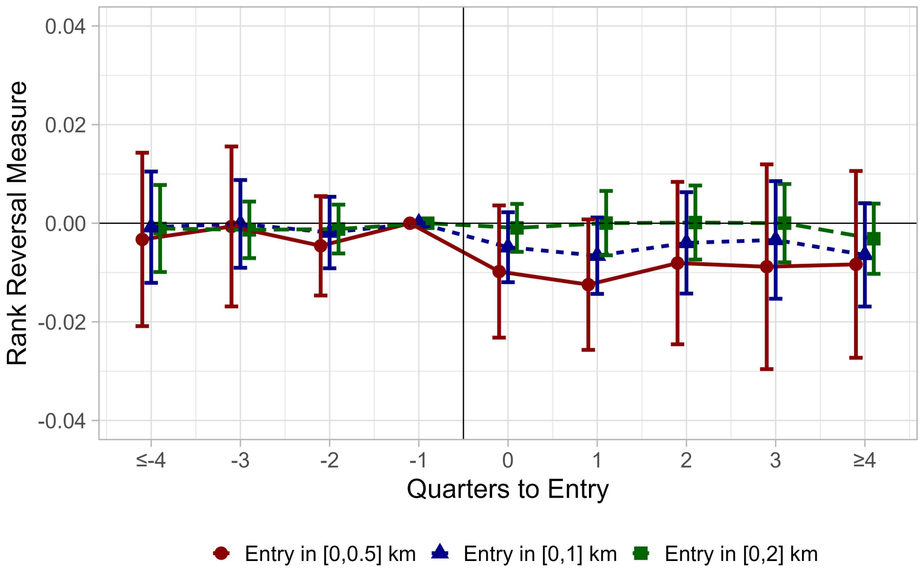
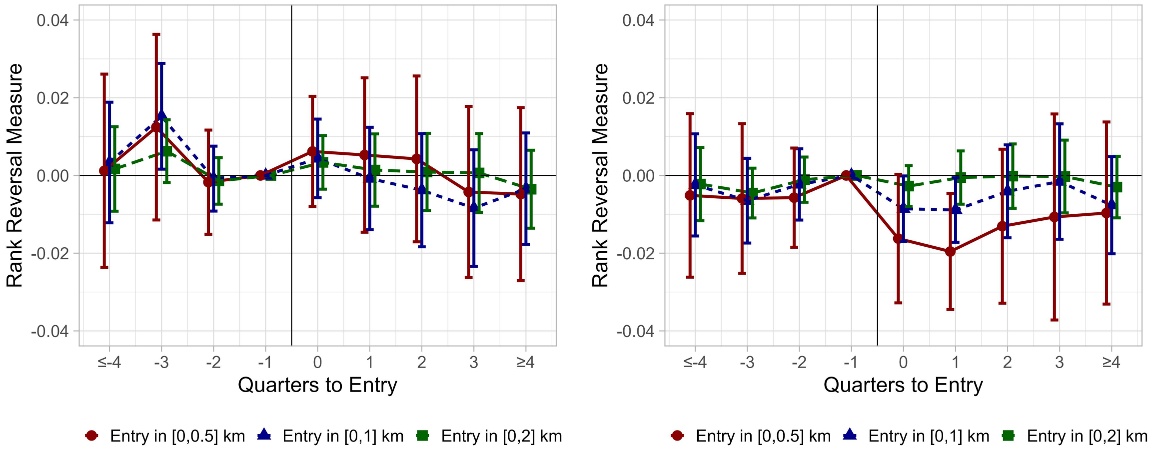


Figure 2.A.1: Entry Effect on Rank Reversal

Note: This figure provides the leads and lags of the event study regression (2.7) with an effect window of four bins before and five bins after the entry event. The outcome variable is the rank reversal measure per couple and quarter constructed as in Chandra and Tappata (2011). The dataset includes all rank reversals of station pairs with a bilateral distance of 2 km maximum. This follows our market definitions from above. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

frictions.

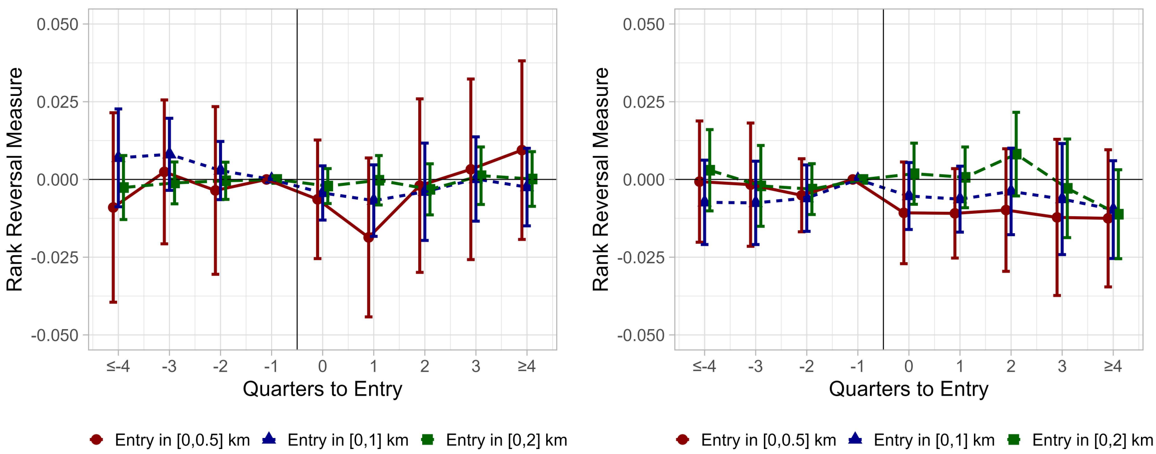


(a) Couple Distance in [0,1] km

(b) Couple Distance in (1,2] km

Figure 2.A.2: Entry Effect on Rank Reversal by Couple Distance

Note: This figure provides the leads and lags of the event study regression (2.7) with an effect window of four bins before and five bins after the entry event. The outcome variable is the rank reversal measure per couple and quarter constructed as in Chandra and Tappata (2011). The dataset includes all rank reversals of station pairs with a bilateral distance of 2 km maximum. This follows our market definitions from above. The left plot looks at couples with a bilateral distance of 1 km maximum. The right plot looks at couples with a bilateral distance of between 1 km and 2 km. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.



(a) Entrant is Nearest Neighbor

(b) Entrant is not Nearest Neighbor

Figure 2.A.3: Entry Effect on Rank Reversal by Entrant Characteristic

Note: This figure provides the leads and lags of the event study regression (2.7) with an effect window of four bins before and five bins after the entry event. The outcome variable is the rank reversal measure per couple and quarter constructed as in Chandra and Tappata (2011). The dataset includes all rank reversals of station pairs with a bilateral distance of 2 km maximum. This follows our market definitions from above. The left plot looks at couples where the entrant is the nearest neighbor. The right plot looks at couples for which this is not the case. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

Appendix B: Additional Figures

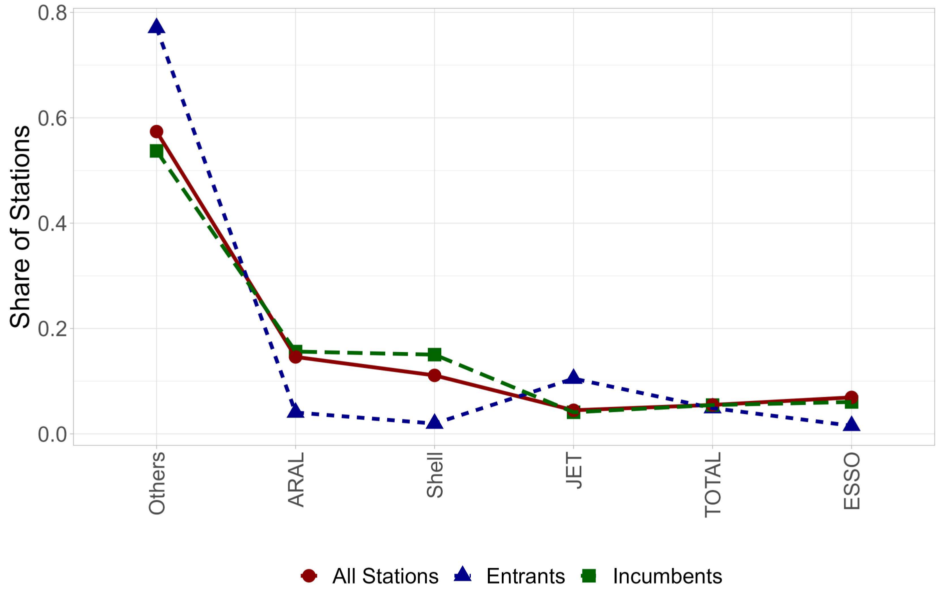


Figure 2.B.1: Brand Distribution

Note: This figure shows the share of stations per brand among all German gasoline stations; for entrants only; and for incumbents only. Incumbents refer to stations in a 1 km radius around entrants. Other brands are pooled in 'Others'.

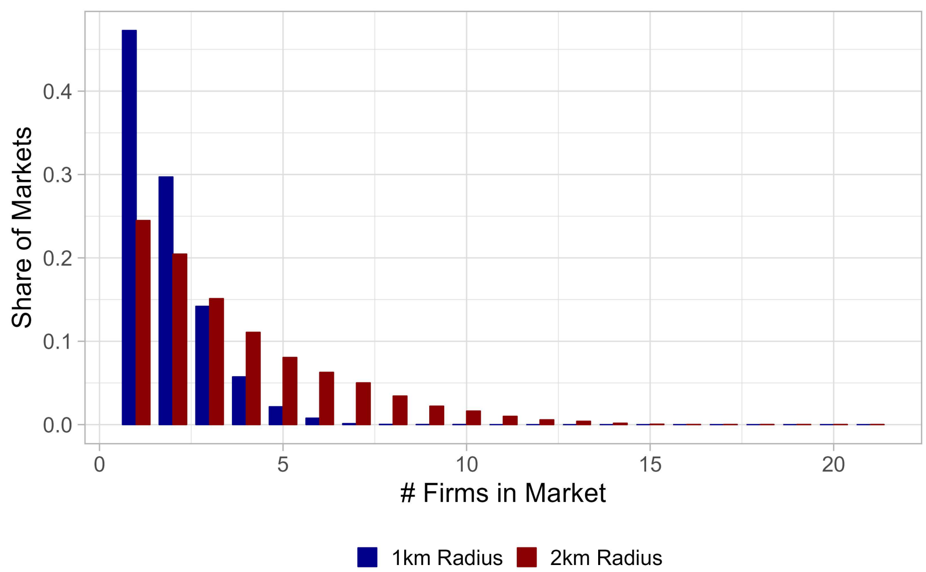


Figure 2.B.2: Market Size Distribution - 1 km and 2 km Radius

Note: The histogram plots the distribution of the number firms in a market using our baseline market definitions: Circles of 1 km or 2 km linear distance around each station. Markets around entrants are not included. Market size is defined at the market-date level.

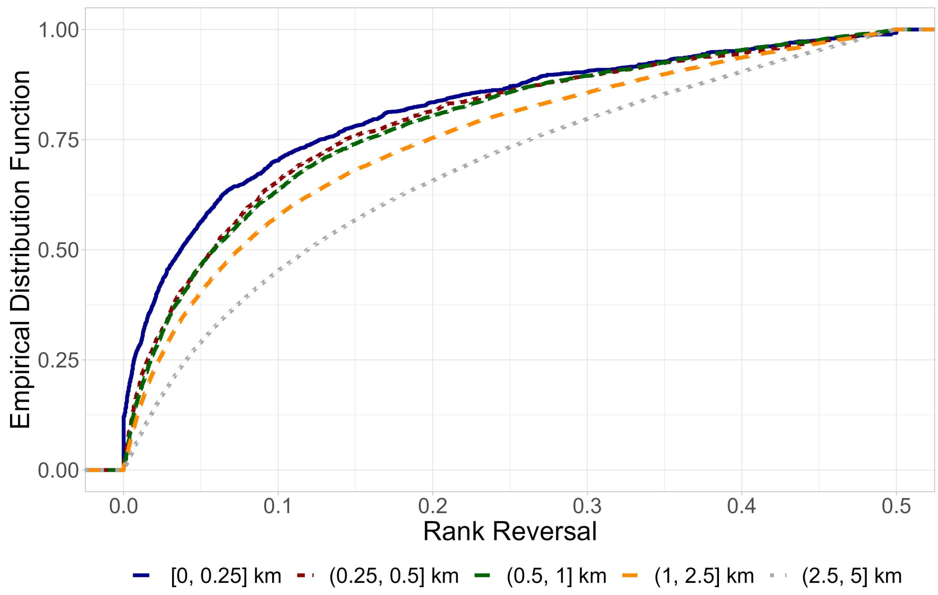


Figure 2.B.3: Rank Reversals - Chandra and Tappata (2011)

Note: Empirical distribution functions of station couples with heterogeneous distances between stations. The rank reversal measure follows the definition in Chandra and Tappata (2011): The share of observations, in which the mostly cheaper station, is the more expensive one.

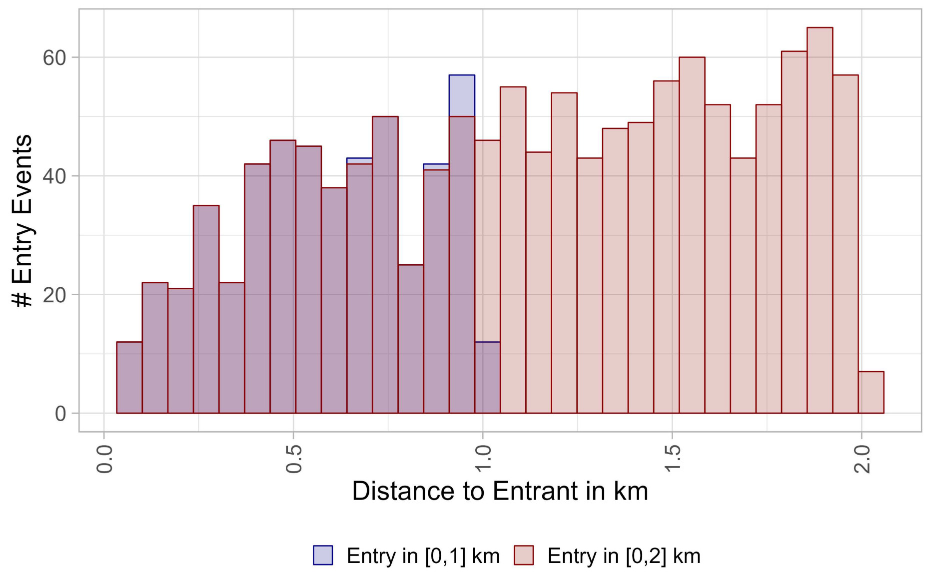
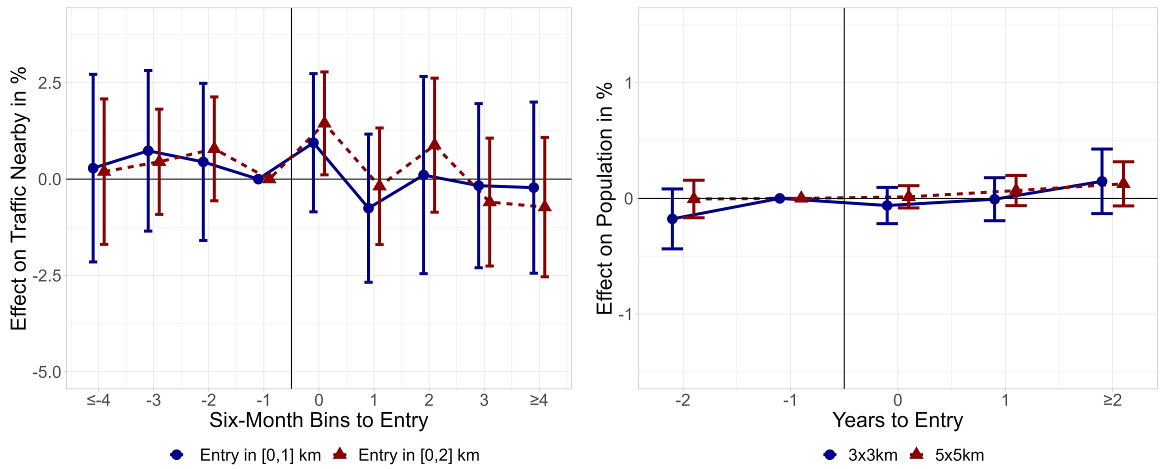


Figure 2.B.4: Distance Between Incumbents and Their Entrants

Note: Small differences between bins of entry in 1 km and 2 km radius for bins between 0 and 1 km distance stem from the fact that only the first entry is used as treatment in our baseline.



(a) Effects on Traffic near Entrants

(b) Effects on Population near Entrants

Figure 2.B.5: Effect of Demand Proxies

Note: The figures plot the effect of station entry on traffic flows between 4 pm and 6 pm at traffic counters in Germany as well as on annual population data in a 3x3km grid around stations. Traffic counters in less than 1 km or 2 km to entry are treated in this analysis. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included. Observations after a second entry are not included in the traffic analysis. In the population analysis, entrants and the population in a 3x3km grid around them are treated, so that no multiple treatment occurs.

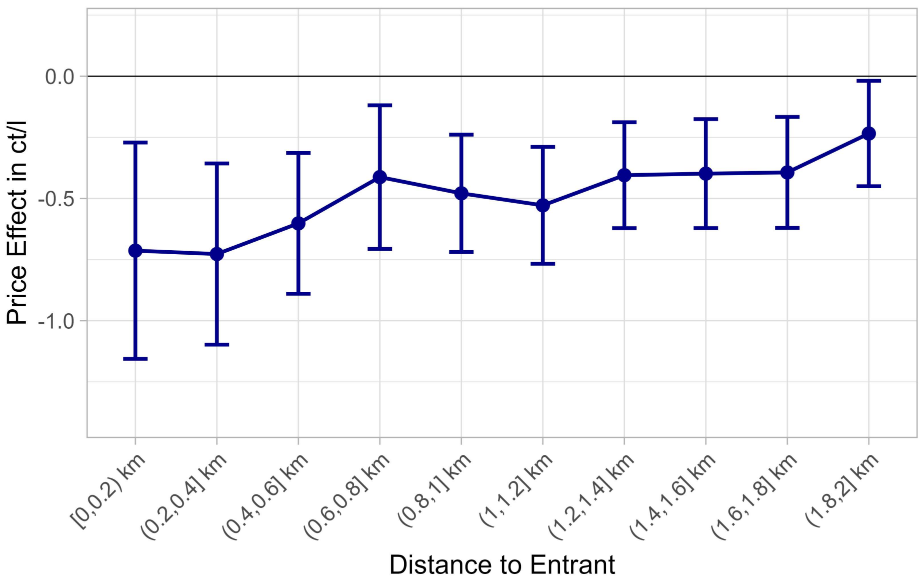


Figure 2.B.6: Effect of Entry on Prices by Distance to Entry

Note: This plot shows the effect of station entry on prices by distance between incumbent and entrant. Only the first entry event for incumbents within 2 km radius is used. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included. Observations after a second entry are not included.

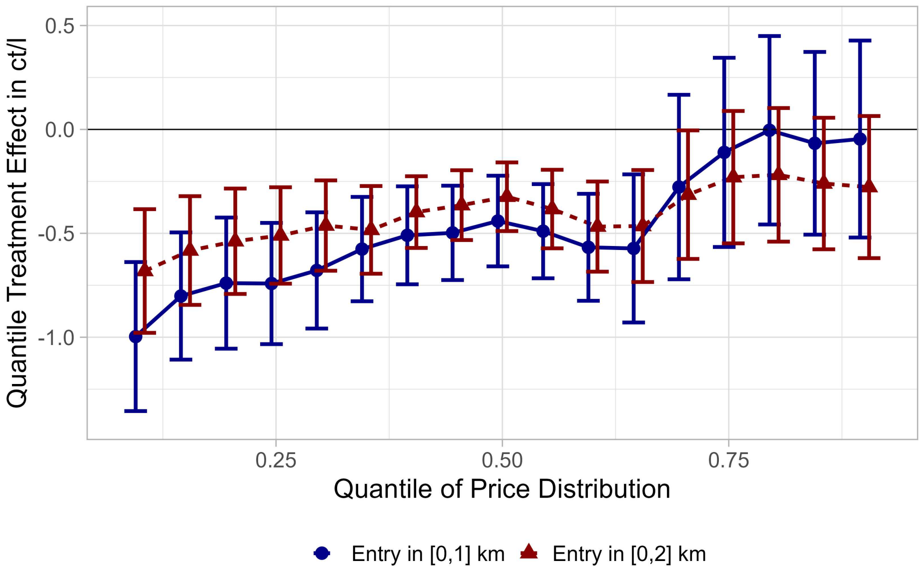
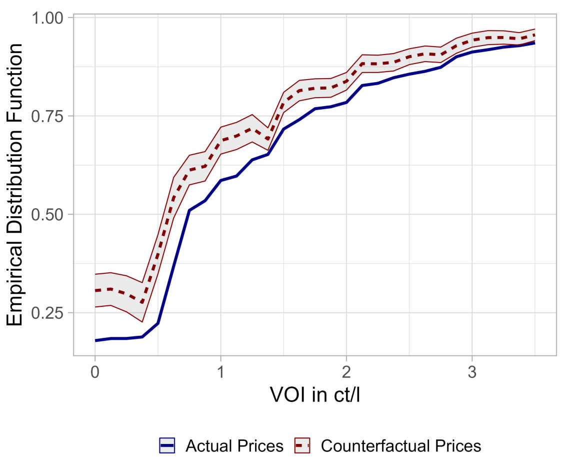
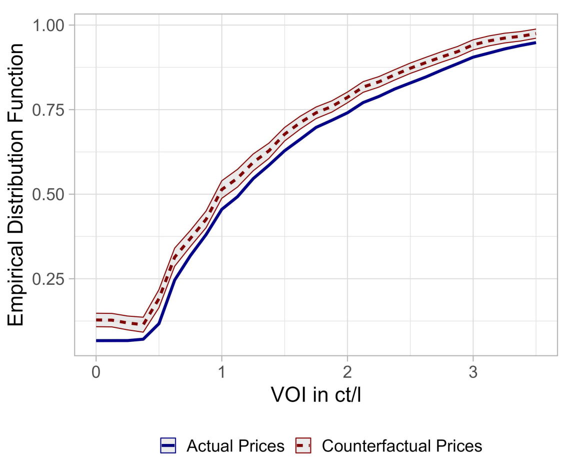


Figure 2.B.7: QTEs of Entry on Prices - Controlling for Wholesale Prices

Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). Estimates for every 5th percentile between the 10th and 90th percentile are provided. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls are included. Results obtained from regressions in which we also control for daily, region-specific wholesale price variations.



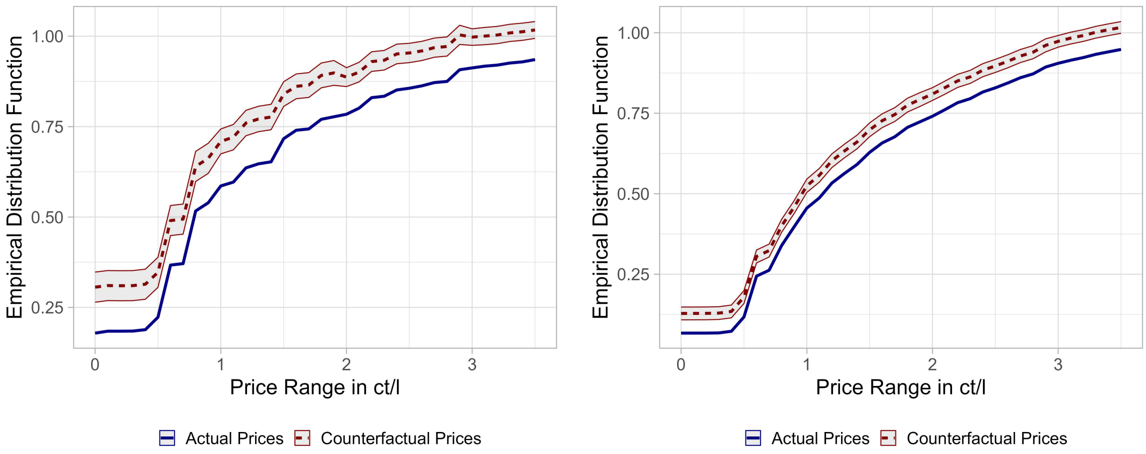
(a) Entry in [0,1] km



(b) Entry in [0,2] km

Figure 2.B.8: Distributional Effects of Dispersion: VOI

Note: The figures plot the actual price dispersion distribution of markets, which experienced entry (blue) and the estimated counterfactual price dispersion distribution (red) of a scenario without entry in the treated markets. The counterfactual distribution comes from distribution regressions as proposed by Chernozhukov et al. (2013), i.e. the regressions as in equation (2.5) for dispersion thresholds in 0.125 ct/l steps. The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. The shaded area indicates the 95% confidence interval.



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.B.9: Distributional Effects of Dispersion: Range

Note: The figures plot the actual price dispersion distribution of markets, which experienced entry (blue) and the estimated counterfactual price dispersion distribution (red) of a scenario without entry in the treated markets. The counterfactual distribution comes from distribution regressions as proposed by Chernozhukov et al. (2013), i.e. the regressions as in equation (2.5) for dispersion thresholds in 0.125 ct/l steps. The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. The shaded area indicates the 95% confidence interval.

Appendix C: Additional Tables

Table 2.C.1: Rank Reversals for Different Station Pair Distances

	rr_{AB}	
	(1)	(2)
1[0 km $\leq$ Distance $\leq$ 0.25 km]	-0.057^{***}	
	(0.006)	
1[0.25 km < Distance $\leq$ 0.5 km]		0.014^{**}
		(0.006)
1[0.5 km < Distance $\leq$ 1 km]		0.016^{***}
		(0.005)
1[1 km < Distance $\leq$ 2.5 km]		0.035^{***}
		(0.005)
1[2.5 km < Distance $\leq$ 5 km]		0.071^{***}
		(0.006)
Observations	81,415	81,415

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are twoway-clustered for the municipality of each station included in the couple. All station pairs of maximum distance 5km are included.

Table 2.C.2: Differences in Levels Between Entrants and Incumbents

	$1[\text{Entrant}]_{it}$					
Market Definition:		[0,1] km			[0,2] km	
Station-specific characteristics.						
Price (ct/l)	−0.0184***	−0.0180***	−0.0083	−0.0085*	−0.0080*	0.0033
	(0.0052)	(0.0052)	(0.0055)	(0.0046)	(0.0046)	(0.0043)
Gross Margin (ct/l)	0.0028	0.0029	0.0100	−0.0067	−0.0063	−0.0040
	(0.0060)	(0.0060)	(0.0065)	(0.0053)	(0.0052)	(0.0049)
# Competitors 1 km Radius	−0.1700***	−0.1607***	−0.1485***			
	(0.0133)	(0.0137)	(0.0118)			
1[# Competitors 1 km > 0]×VOI		−0.0248***	−0.0114***			
		(0.0054)	(0.0058)			
# Competitors 2 km Radius				−0.0420***	−0.0391***	−0.0384***
				(0.0066)	(0.0065)	(0.0060)
1[# Competitors 2 km > 0]×VOI					−0.0208***	−0.0143***
					(0.0065)	(0.0049)
log(Population - 3x3 km Grid)	−0.0779***	−0.0759***	−0.0646***	−0.1291***	−0.1251***	−0.0997***
	(0.0127)	(0.0126)	(0.0125)	(0.0131)	(0.0130)	(0.0122)
County-specific characteristics.						
log(Unemployment Rate)	0.0084	0.0070	0.0630	−0.0225	−0.0198	0.0192
	(0.0450)	(0.0445)	(0.0510)	(0.0378)	(0.0371)	(0.0419)
log(Commuters)	0.0316	0.0302	0.0524	−0.1088	−0.1028	−0.0522
	(0.0774)	(0.0760)	(0.0850)	(0.0673)	(0.0658)	(0.0677)
log(Population)	0.1787*	0.1788*	0.0705	0.3414***	0.3349***	0.2635***
	(0.0955)	(0.0942)	(0.1065)	(0.0860)	(0.0851)	(0.0939)
log(Vehicles)	−0.2112*	−0.2096*	−0.1341	−0.2153**	−0.2153**	−0.1910*
	(0.1146)	(0.1136)	(0.1276)	(0.1024)	(0.1021)	(0.1120)
log(GDP p.c.)	0.0238	0.0256	0.0213	0.0511	0.0522	0.0356
	(0.0412)	(0.0406)	(0.0422)	(0.0376)	(0.0374)	(0.0406)
log(Available Income p.c.)	−0.0407	−0.0302	0.0532	0.0229	0.0325	0.1017
	(0.1205)	(0.1184)	(0.1350)	(0.1174)	(0.1164)	(0.1198)
Brand FE	×	×	✓	×	×	✓
State-Date FE	✓	✓	✓	✓	✓	✓
Observations		1,103,982			1,838,558	

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the municipality level. Linear probability model with dichotomous outcome. Only observations of entrants and incumbents after entry are included.

Table 2.C.3: Differences in Levels Between Incumbents and Outsiders

	$1[\text{Incumbent}]_{it}$					
Market Defintion:		[0,1] km			[0,2] km	
Station-specific characteristics.						
Price (ct/l)	0.0008	0.0008	0.0006	0.0005	0.0005	0.0002
	(0.0006)	(0.0006)	(0.0006)	(0.0013)	(0.0014)	(0.0013)
Gross Margin (ct/l)	−0.0002	−0.0002	−0.0000	0.0007	0.0007	0.0011
	(0.0007)	(0.0007)	(0.0007)	(0.0015)	(0.0015)	(0.0015)
# Competitors 1 km Radius	0.0005	0.0004	0.0005			
	(0.0010)	(0.0011)	(0.0011)			
1[# Competitors 1 km > 0]×VOI		0.0001	−0.0001			
		(0.0008)	(0.0008)			
# Competitors 2 km Radius				0.0025	0.0024	0.0023
				(0.0017)	(0.0018)	(0.0018)
1[# Competitors 2 km > 0]×VOI					0.0010	0.0005
					(0.0014)	(0.0013)
log(Population - 3x3 km Grid)	0.0029***	0.0029***	0.0027***	0.0097***	0.0096***	0.0092***
	(0.0010)	(0.0010)	(0.0011)	(0.0022)	(0.0022)	(0.0022)
County-specific characteristics.						
log(Unemployment Rate)	−0.0059	−0.0059	−0.0070	−0.0130	−0.0130	−0.0156
	(0.0065)	(0.0066)	(0.0064)	(0.0146)	(0.0146)	(0.0145)
log(Commuters)	0.0038	0.0038	0.0030	0.0075	0.0076	0.0066
	(0.0095)	(0.0095)	(0.0096)	(0.0247)	(0.0246)	(0.0244)
log(Population)	−0.0118	−0.0118	−0.0103	−0.0316	−0.0317	−0.0283
	(0.0114)	(0.0114)	(0.0115)	(0.0286)	(0.0286)	(0.0284)
log(Vehicles)	0.0064	0.0064	0.0055	0.0139	0.0139	0.0110
	(0.0135)	(0.0135)	(0.0137)	(0.0346)	(0.0346)	(0.0343)
log(GDP p.c.)	0.0008	0.0008	0.0007	0.0004	0.0004	0.0007
	(0.0052)	(0.0052)	(0.0052)	(0.0128)	(0.0128)	(0.0128)
log(Available Income p.c.)	−0.0290*	−0.0289*	−0.0286*	−0.0715*	−0.0715*	−0.0709*
	(0.0153)	(0.0153)	(0.0149)	(0.0390)	(0.0390)	(0.0384)
Brand FE	×	×	✓	×	×	✓
State-Date FE	✓	✓	✓	✓	✓	✓
Observations		24,411,307			23,676,731	

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the municipality level. Linear probability model with dichotomous outcome. Observations of entrants are dropped. Post-treatment observations of incumbents are dropped.

Table 2.C.4: (Average) Effect of Entry on Prices

	$Price_{it}$	
	Entry in [0,1] km	Entry in [0,2] km
$1[Post - Entry]_{it}$	-0.536^{***}	-0.450^{***}
	(0.084)	(0.063)
Station FE	✓	✓
State-Date FE	✓	✓
Observations	24,887,154	24,860,948

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the municipality level. County-level controls are included. Observation numbers differ across treatment allocation as observations of stations, which experience entry more than once, are dropped from the sample after the date of the second entry.

Appendix D: Online Appendix

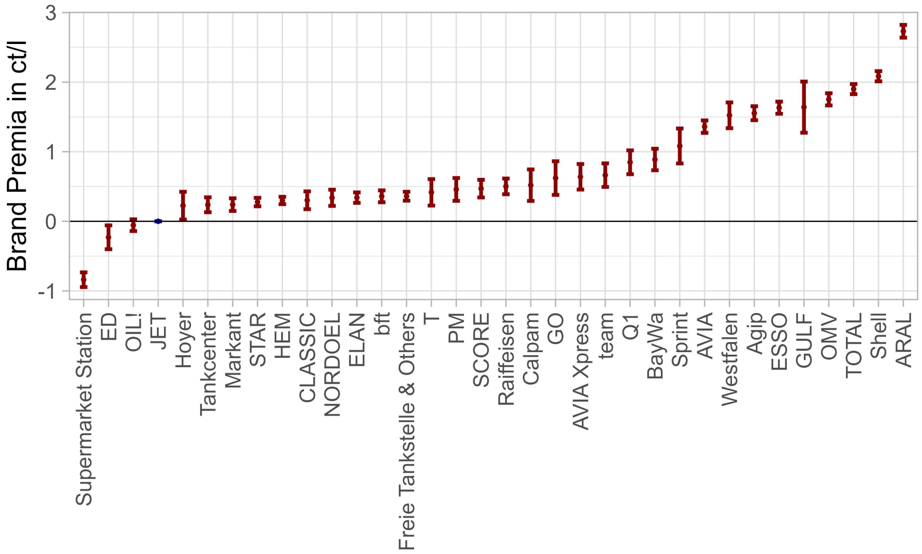
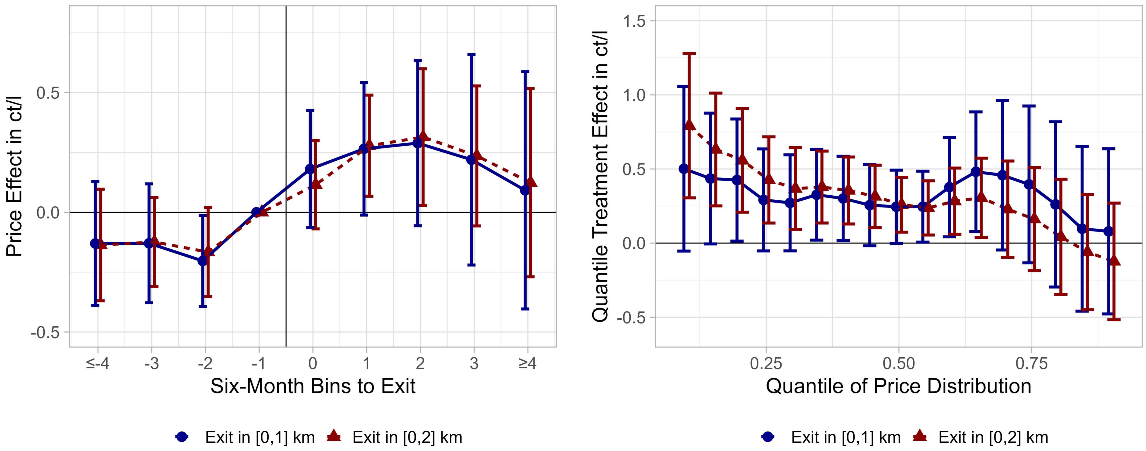


Figure 2.D.1: Brand-Specific Price Premia

Note: Coefficients obtained from a regression of prices on brand fixed effects, state-date fixed effects. Standard errors are clustered at municipality level. Vertical bars indicate 95% confidence intervals.



(a) Average Price Effect (b) QTE - Exit

Figure 2.D.2: Robustness Check: Exit - Average Price Effect and QTEs

Note: The left figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the exit event. The right figure plots quantile treatment effects of exit on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009), where the effects give the coefficient of all lags binned together. We present estimates for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

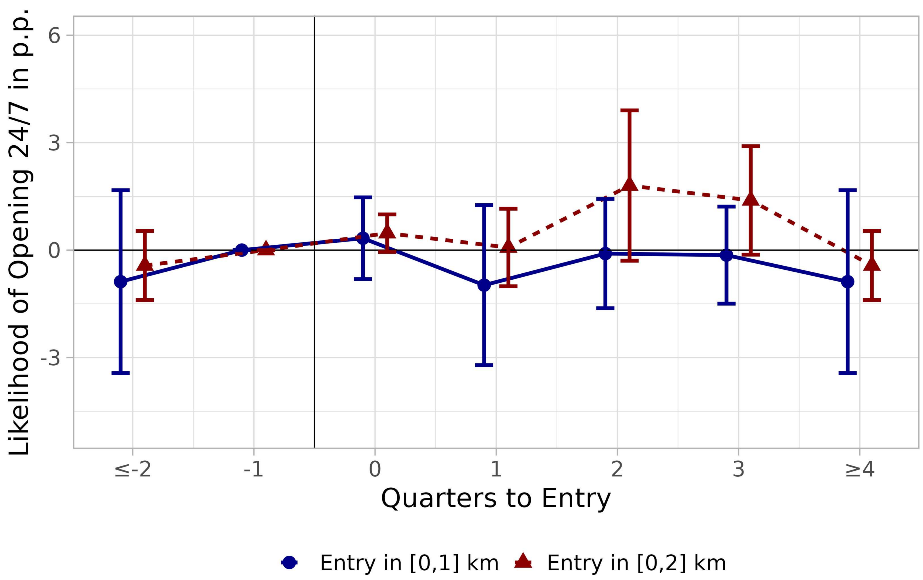
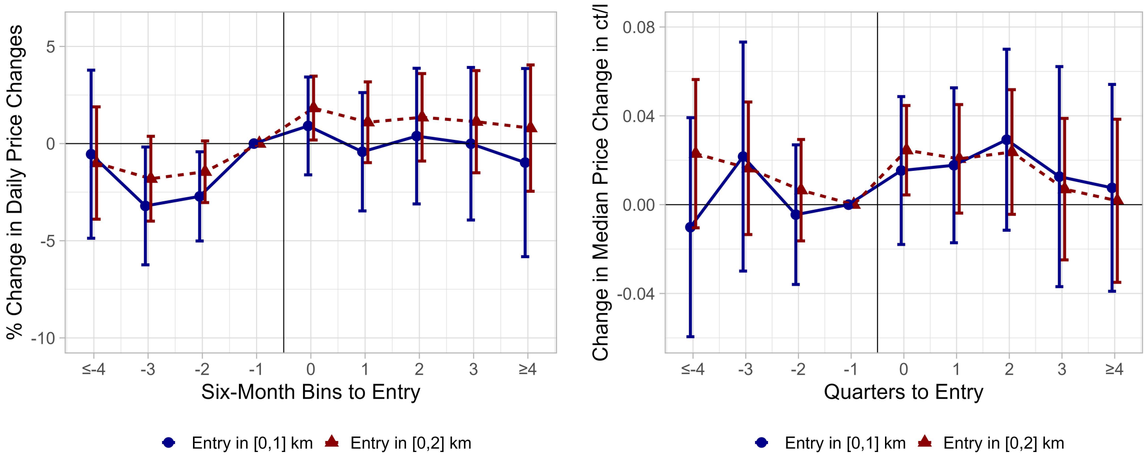


Figure 2.D.3: Robustness Check: Effect of Entry on Opening Hours

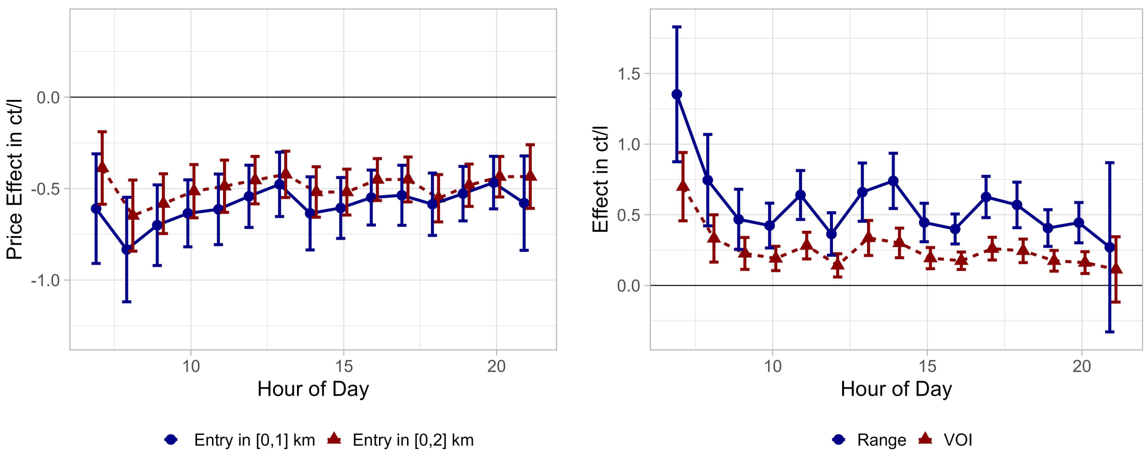
Note: This figure provides the leads and lags, β_τ, of the event study regression (2.3) with a shorter effect window and thinner (quarterly) bins as opening hours are observed for the period January 2019 to March 2020. The outcome is a dummy equal to one if a station is open 24/7, and zero otherwise. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.



(a) Price Changes (b) Median Price Change

Figure 2.D.4: Robustness Check: Price Change Characteristics

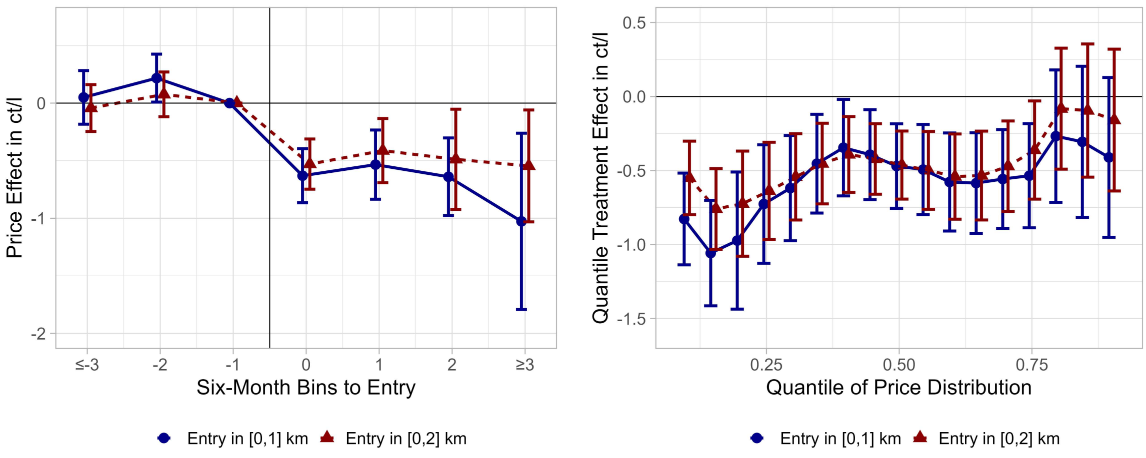
Note: This figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.



(a) Average Price Effect (b) Av. Dispersion Effect - Entry in [0,1] km

Figure 2.D.5: Robustness Check: Effect of Entry by Hour of Day

Note: This figure provides the estimates of the baseline difference-in-differences regression (2.2) with prices and market-level dispersion as the outcome. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.

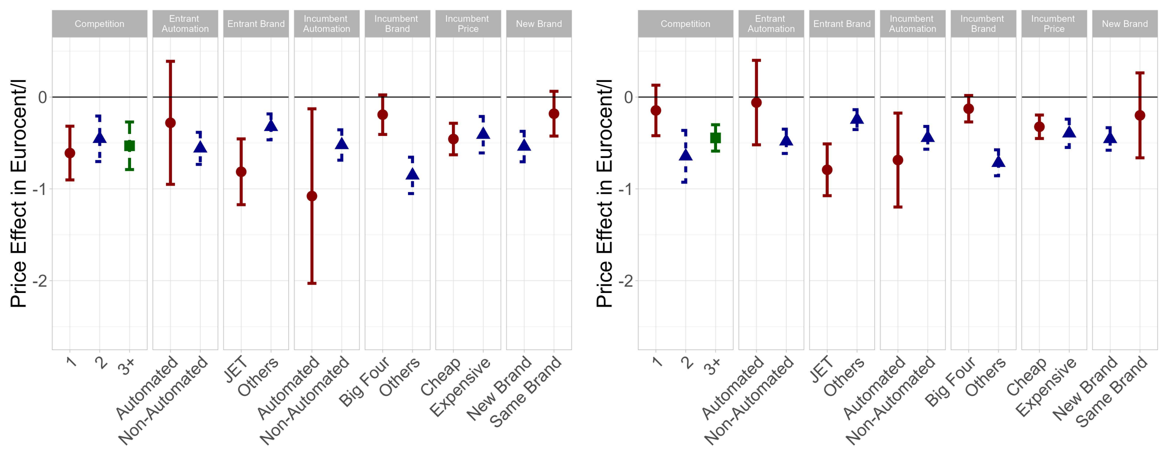


(a) Average Price Effect

(b) Quantile Treatment Effects

Figure 2.D.6: Robustness Check: Pre-Algorithm Adoption Time Period

Note: The left figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. The right figure plots quantile treatment effects of entry on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). We present estimates for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.7: Robustness Check: Heterogeneity in (Average) Effect of Entry on Prices

Note: The plots give simple difference-in-differences results as in equation (2.2), where the treatment variable is interacted with the respective heterogeneity dummy. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.

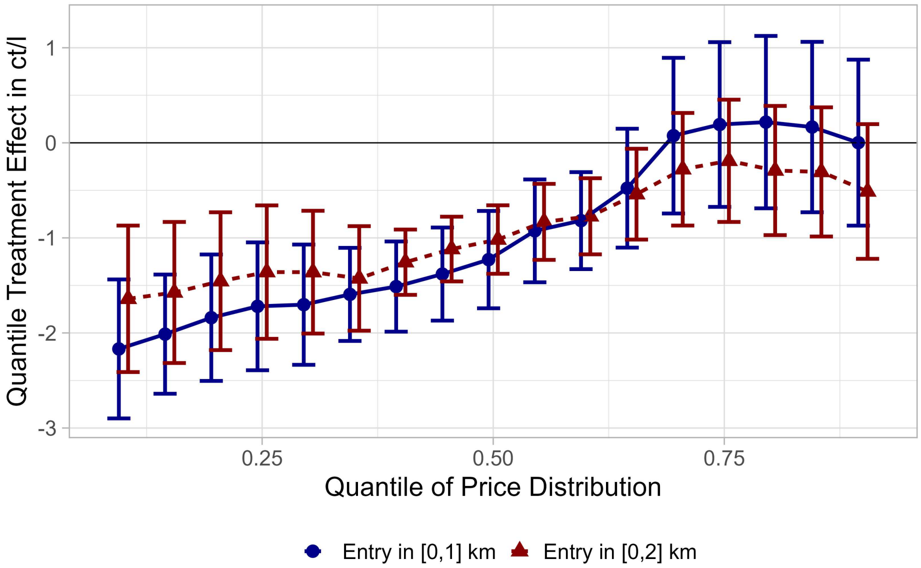
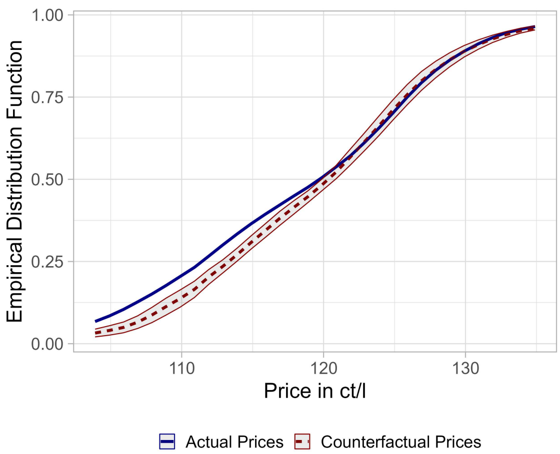
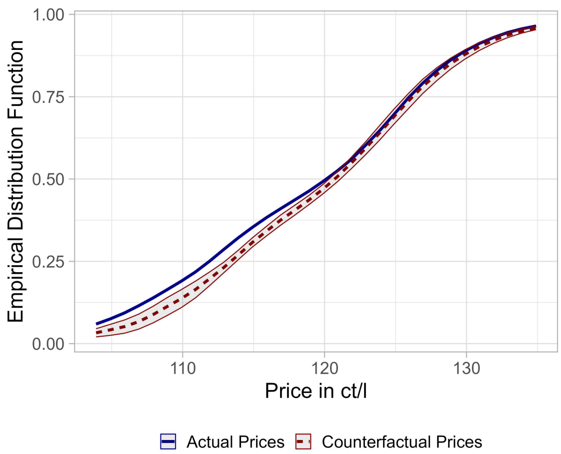


Figure 2.D.8: Robustness Check: JET Entry - QTEs

Note: This figure plots quantile treatment effects of entry by JET on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). Estimates are given for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.



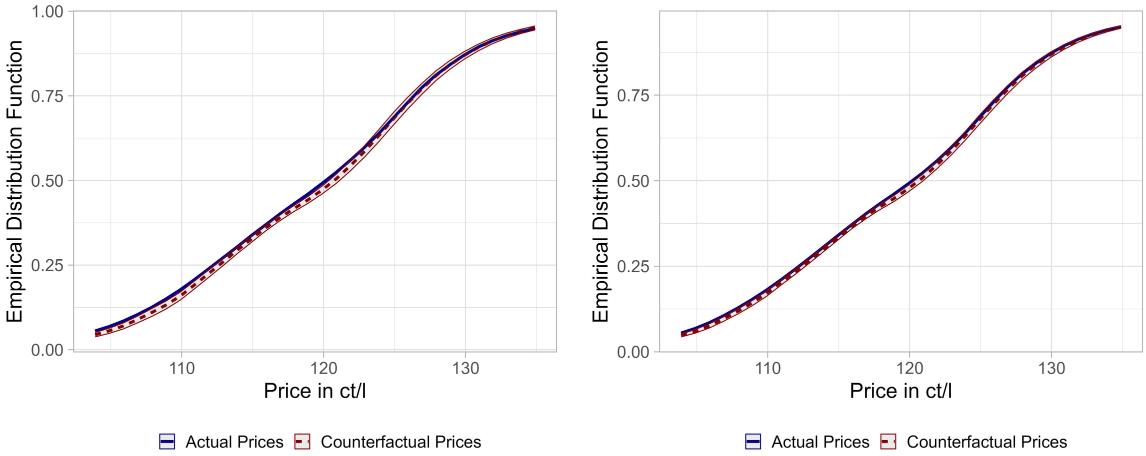
(a) Entry in [0,1] km



(b) Entry in [0,2] km

Figure 2.D.9: Robustness Check: JET Entry - Counterfactual Distribution

Note: The figures plot the actual price distribution of incumbents' prices, which experienced entry by JET (blue) and the estimated counterfactual price distribution (red) of a scenario without entry in the treated markets. The counterfactual distribution comes from distribution regressions as proposed by Chernozhukov et al. (2013), i.e., the regressions as in equation (2.5) for price thresholds at each integer ct/l. The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. The shaded area indicates 95% confidence interval.



(a) Entry in [0,1] km (b) Entry in [0,2] km

Figure 2.D.10: Robustness Check: Non-JET Entry - Counterfactual Distribution

Note: The figures plot the actual price distribution of incumbents' prices, which experienced entry by non-JET brands (blue) and the estimated counterfactual price distribution (red) of a scenario without entry in the treated markets. The counterfactual distribution comes from distribution regressions as proposed by Chernozhukov et al. (2013), i.e., the regressions as in equation (2.5) for price thresholds at each integer ct/l. The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. The shaded area indicates the 95% confidence interval.

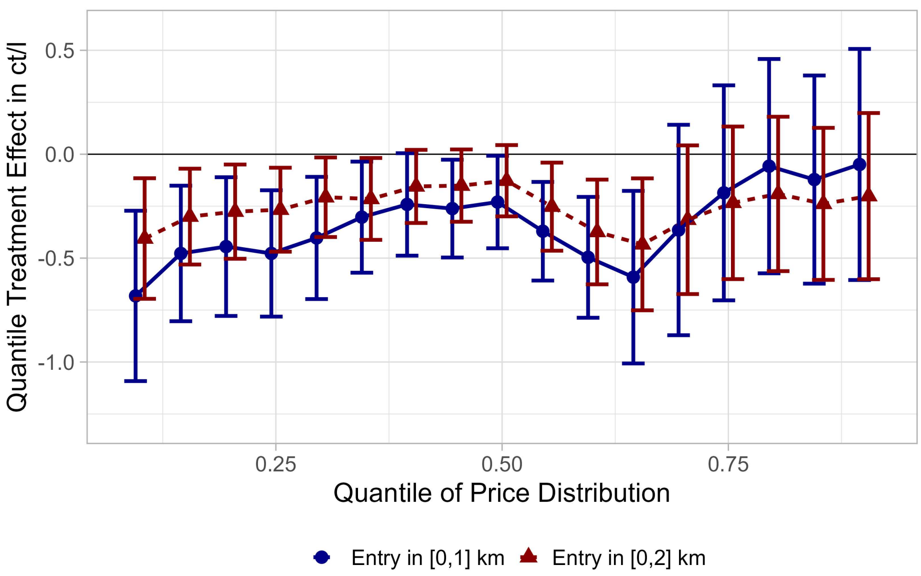
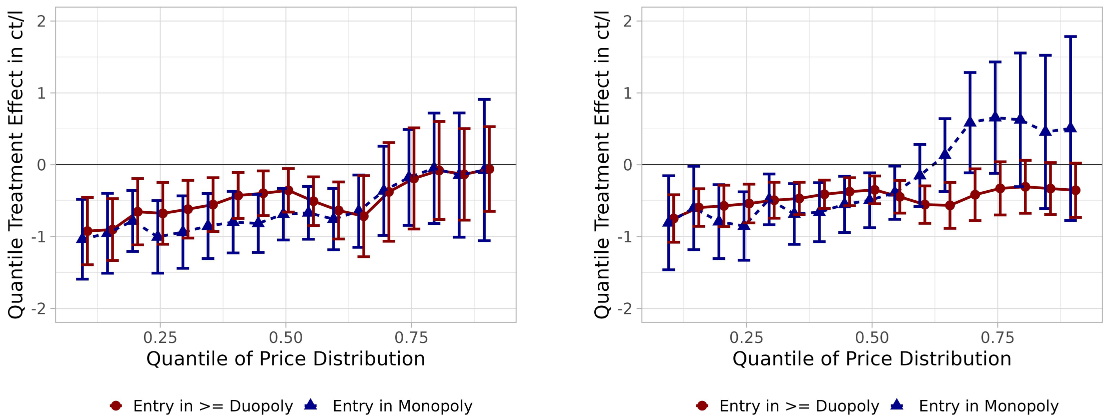


Figure 2.D.11: Robustness Check: Non-JET Entry - QTEs

Note: This figure plots quantile treatment effects of entry by non-JET on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). Estimates are given for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.



(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.12: Robustness Check: Distributional Effects and Market Concentration

Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius and 2 km radius estimated using the method of Firpo et al. (2009). The plots differentiate between markets with few (blue) or many (red) non-entrant stations. As the number of firms in the market increases for a wider market definition, we differently split the sample for comparison reasons. Estimates for every 5th percentile between the 10th and 90th percentile are provided. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.

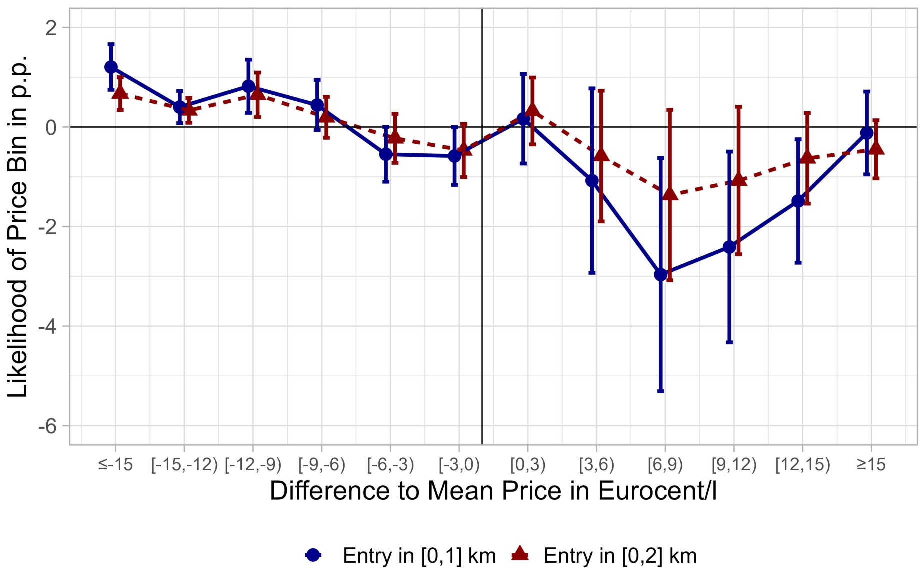
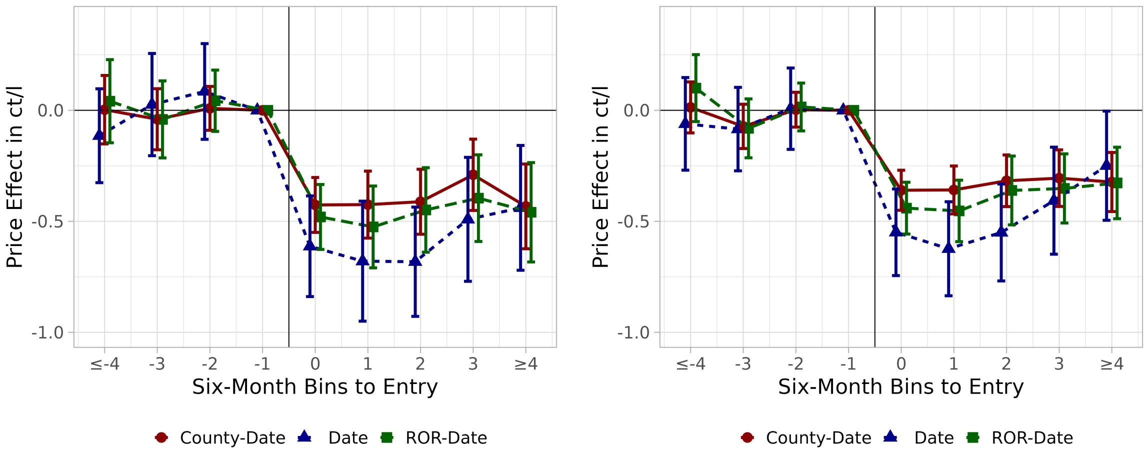


Figure 2.D.13: Robustness Check: Distributional Analysis as in Cengiz et al. (2019)

Note: This figure plots the results of the distributional analysis in the spirit of Cengiz et al. (2019). Results come from difference-in-differences regressions in the style of equation (2.2). Outcome variables for each regression are dummies whether prices are in a specific price bin. We chose price bins of three ct/l width. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.

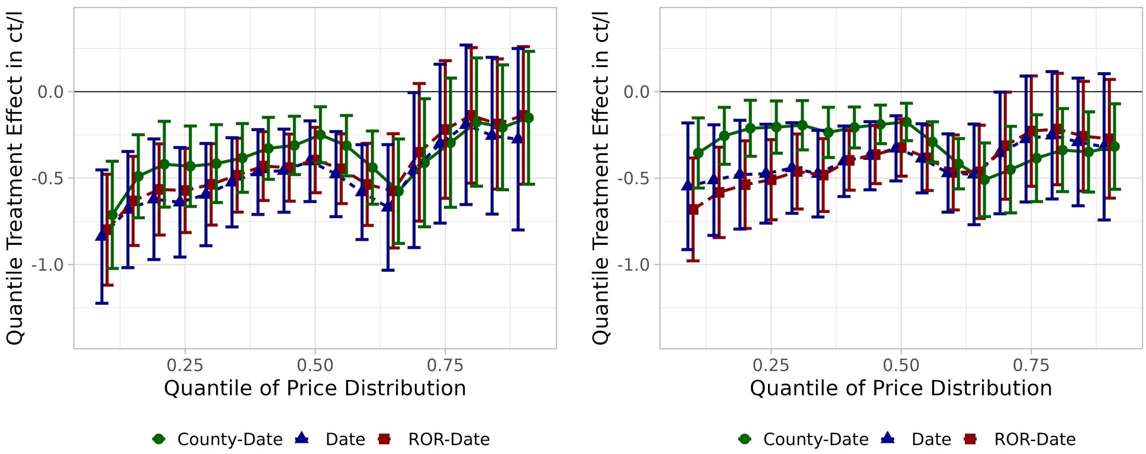


(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.14: Robustness Check: Identification Cells

Note: This figure gives the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Fixed effects are varied. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

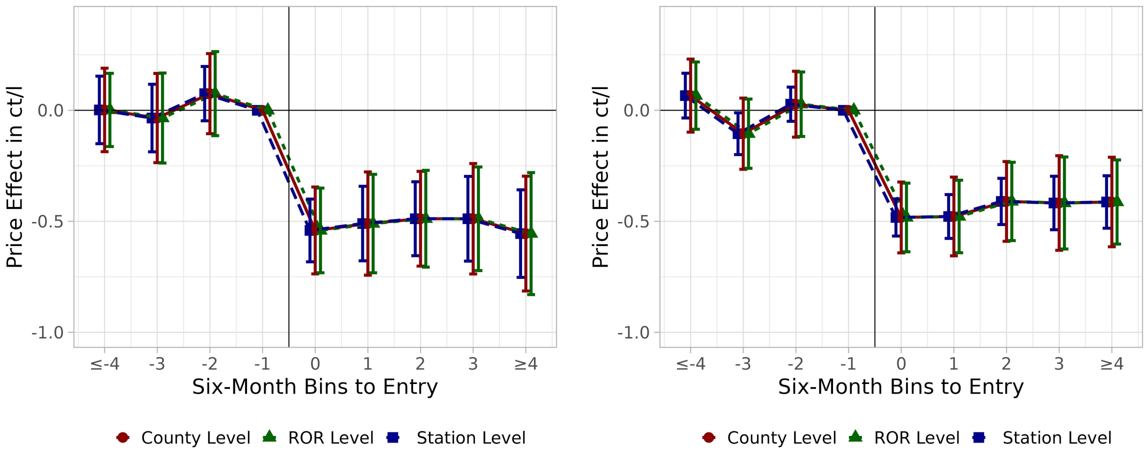


(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.15: Robustness Check: Identification Cells - QTEs

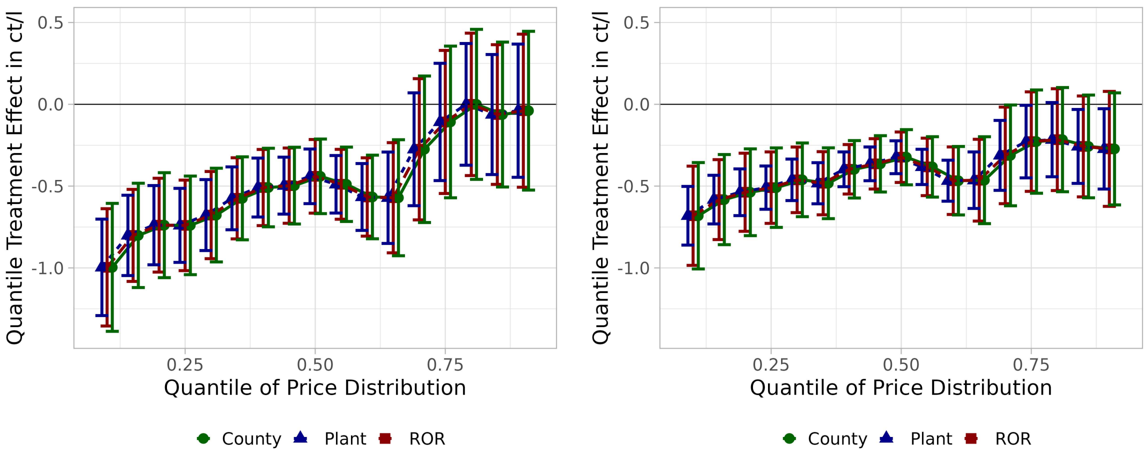
Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius (left) and 2 km radius (right) estimated using the method of Firpo et al. (2009). Estimates are given for every 5th percentile between the 10th and 90th percentile. Fixed effects are varied. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.



(a) Entry in [0,1] km (b) Entry in [0,2] km

Figure 2.D.16: Robustness Check: Inference

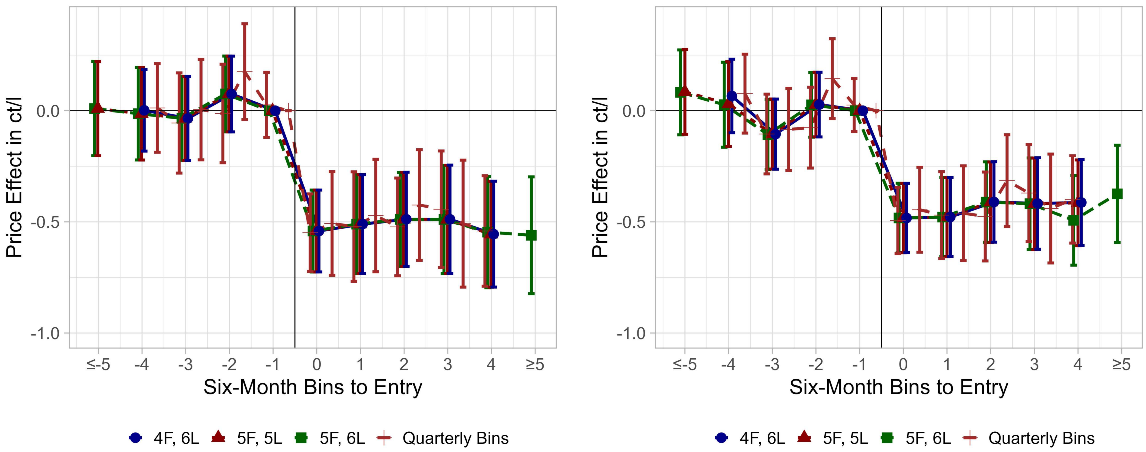
Note: This figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Inference approaches are varied. Standard errors are clustered at different aggregation levels. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.



(a) Entry in [0,1] km (b) Entry in [0,2] km

Figure 2.D.17: Robustness Check: Inference - QTEs

Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius (left) and 2 km radius (right) estimated using the method of Firpo et al. (2009). We present estimates for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at different aggregation levels. Vertical bars indicate 95% confidence intervals. County-level controls included.

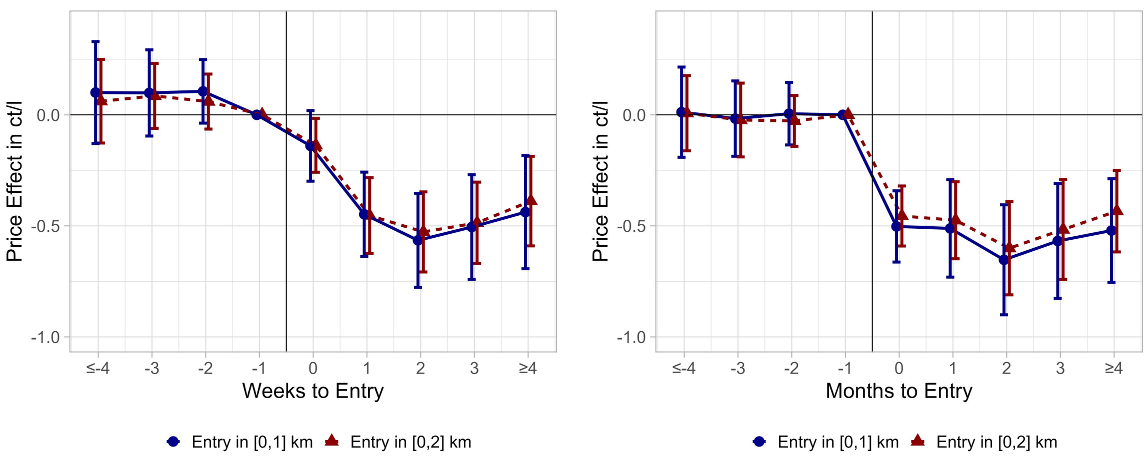


(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.18: Robustness Check: Choice of Leads and Lags

Note: This figure provides the leads and lags of the event study regression (2.3) with a varying effect window. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.

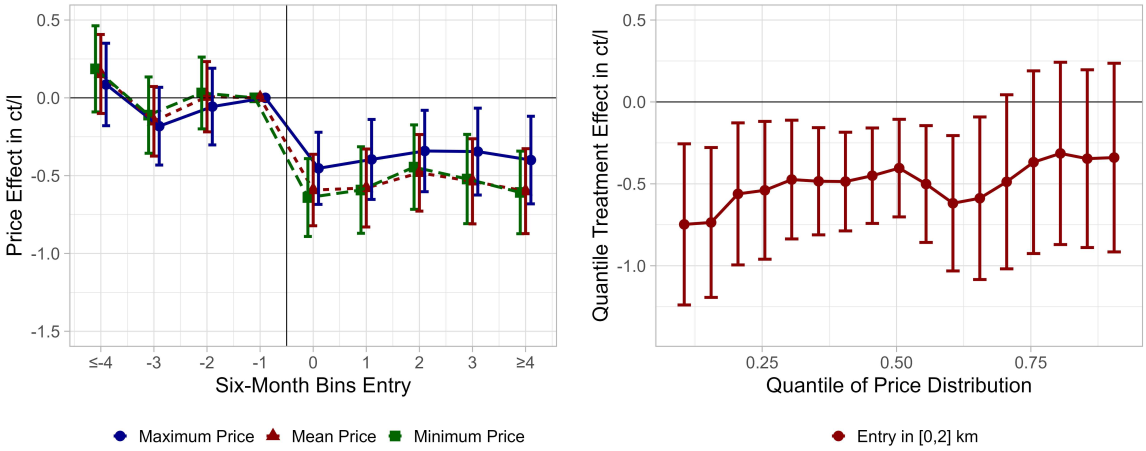


(a) Weekly Bins

(b) Monthly Bins

Figure 2.D.19: Robustness Check: Short-Run Effects

Note: This figure provides the leads and lags of the event study regression (2.3) with a varying effect window. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.

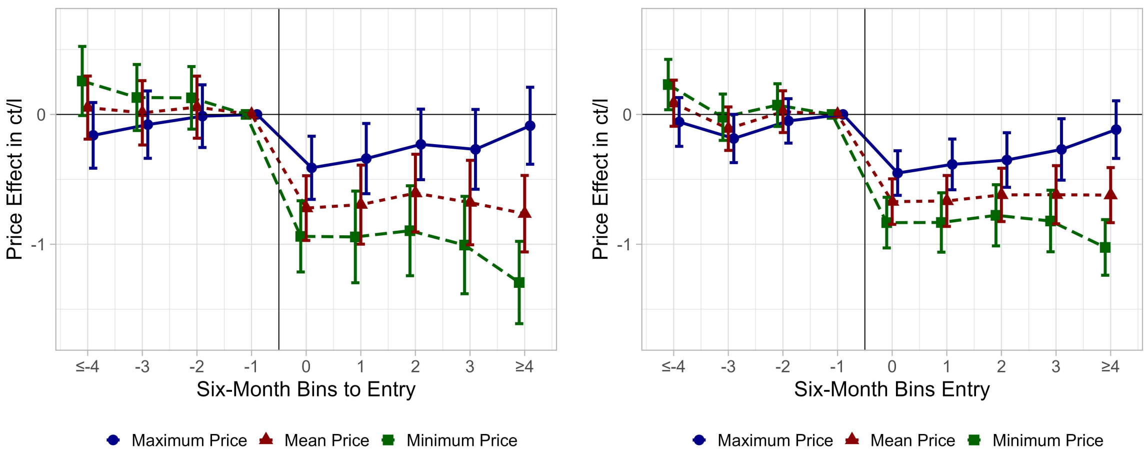


(a) Min, Mean, Max Price

(b) QTEs

Figure 2.D.20: Effect of Entry with $N \geq 4$ Stations

Note: The left figure plots the effect of entry in markets with $N \geq 4$ and 2 km radius on the minimum, mean and maximum price. The right figure plots quantile treatment effects of entry in markets with $N \geq 4$ and 2 km radius estimated using the method of Firpo et al. (2009). Estimates are given for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.

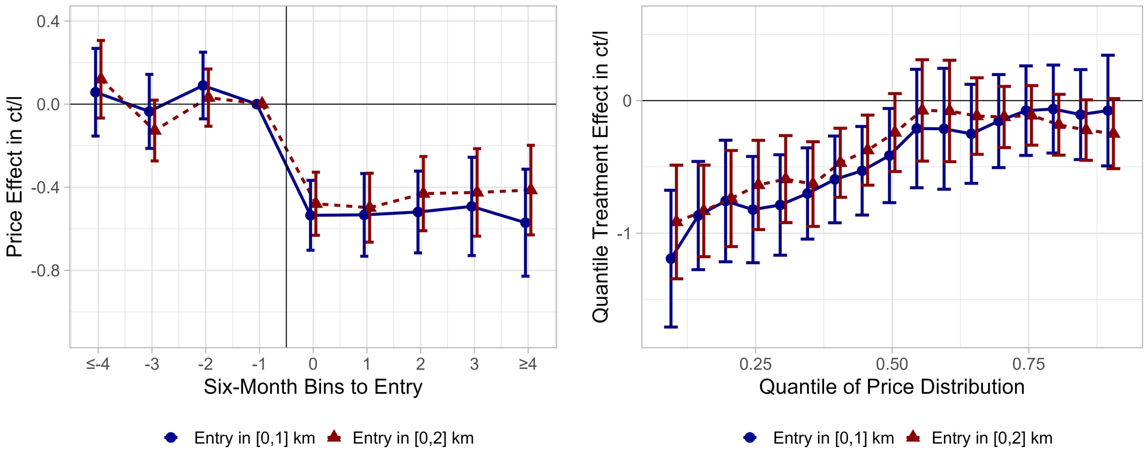


(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.21: Effect of Entry on Price Distribution (Based on Pre-Entry Characteristics)

Note: This figure presents the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Outcome variables are the market-level minimum, mean and maximum price for each market-date observation. Only observations with at least two firms in the market included. Entrants' prices are included in the calculation. Only treated markets with at least two stations in the last pre-entry quarter included. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

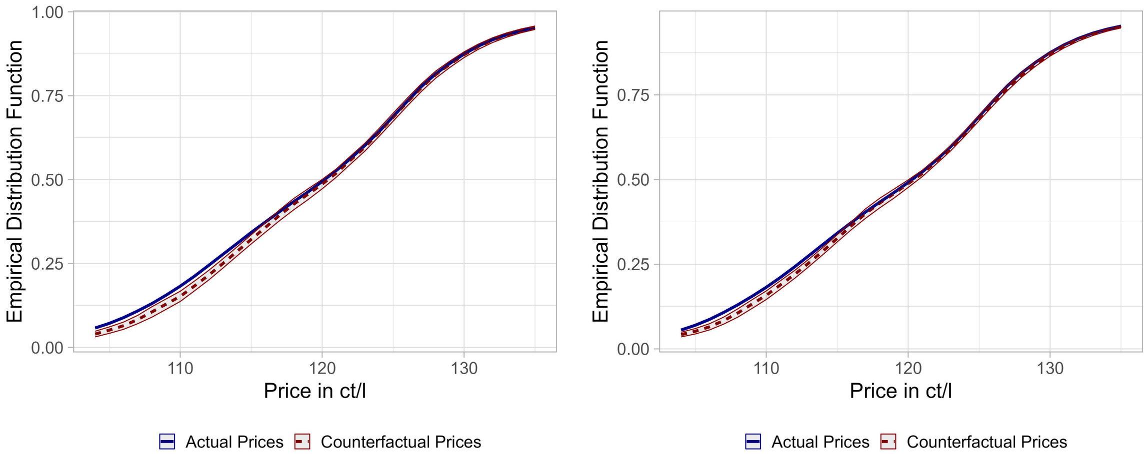


(a) Sun and Abraham (2021) - Event Study

(b) Sun and Abraham (2021) - QTEs

Figure 2.D.22: Robustness Check: Heterogeneous Treatment Effects

Note: The left figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. The right figure plots quantile treatment effects of entry on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009). We present estimates for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.

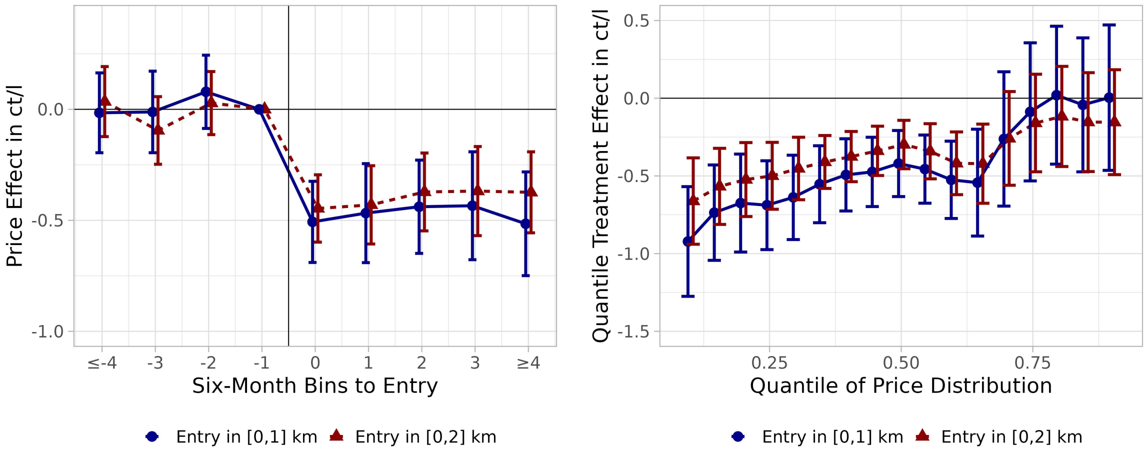


(a) Entry in [0,1] km

(b) Entry in [0,2] km

Figure 2.D.23: Robustness Check: Observed Prices and Counterfactual Distribution - Sun and Abraham (2021)

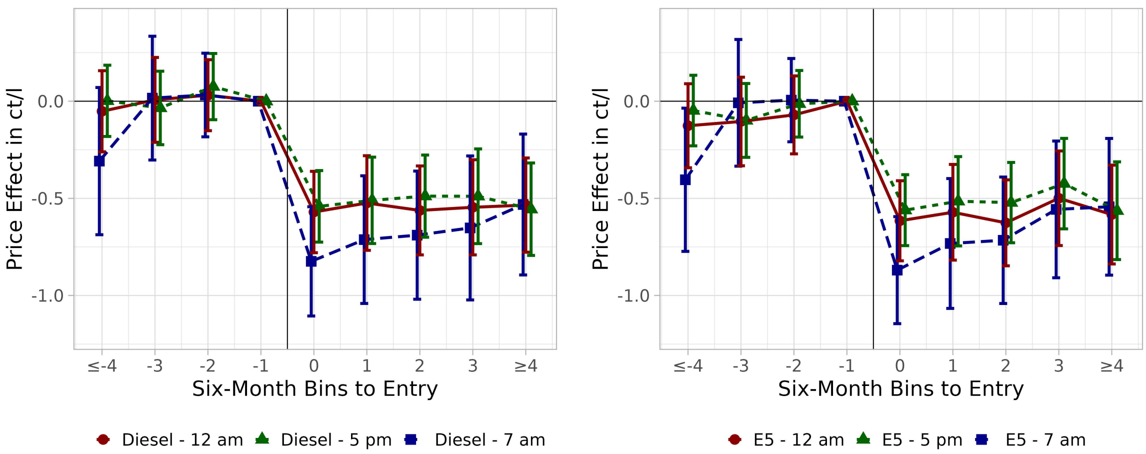
Note: The figures plot the actual price distribution of incumbents' prices, which experienced entry (blue) and the estimated counterfactual price distribution (red) of a scenario without entry in the treated markets. The counterfactual distribution comes from distribution regressions as proposed by Chernozhukov et al. (2013), i.e., the regressions as in equation (2.5) for price thresholds at each integer ct/l. We employ the estimator proposed by Sun and Abraham (2021). The estimated treatment effect per distribution regression is then added to the quantile of the empirical distribution function of actual prices at the respective threshold price. The figure is truncated at the 5th and 95th percentile of the actual price distribution. Standard errors are clustered at the municipality level. The shaded area indicates the 95% confidence interval.



(a) Average Price Effect (b) Quantile Treatment Effects

Figure 2.D.24: Robustness Check: Allowing for Multiple Entries per Incumbent

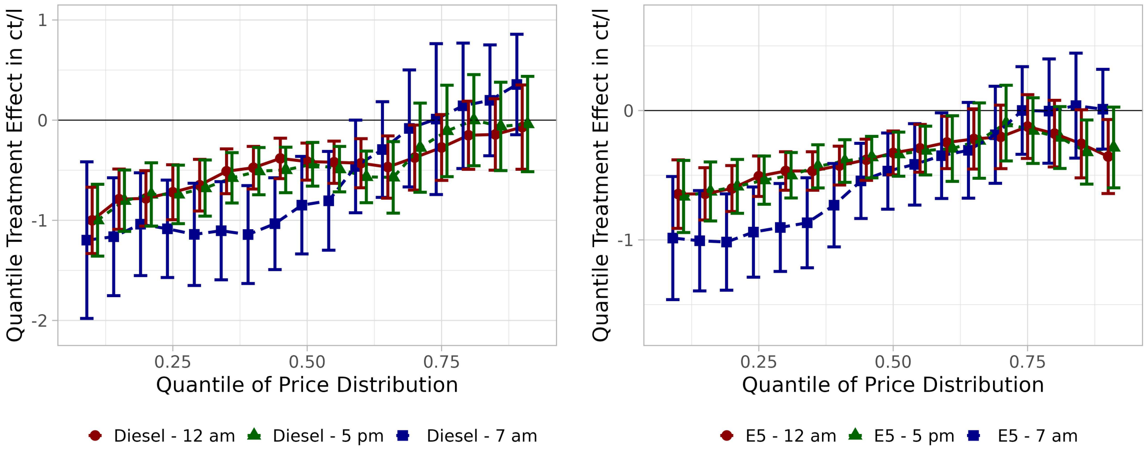
Note: The left figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. We now allow for multiple treatment, so that the treatment is not absorbing. The right figure plots quantile treatment effects of entry on prices for entry in 1 km radius (blue) and 2 km radius (red) estimated using the method of Firpo et al. (2009), where the effects give the coefficient of all lags binned together. We present estimates for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls are included.



(a) Entry in [0,1] km - Diesel (b) Entry in [0,1] km - E5

Figure 2.D.25: Robustness Check: Type of Fuel and Time of the Day

Note: This figure provides the leads and lags of the event study regression (2.3) with an effect window of four bins before and five bins after the entry event. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. Endpoints are binned. County-level controls included.



(a) Entry in [0,1] km - Diesel

(b) Entry in [0,1] km - E5

Figure 2.D.26: Robustness Check: Fuel and Time of the Day - QTEs

Note: This figure plots quantile treatment effects of entry on prices for entry in 1 km radius estimated using the method of Firpo et al. (2009). Estimates are given for every 5th percentile between the 10th and 90th percentile. Standard errors are clustered at the municipality level. Vertical bars indicate 95% confidence intervals. County-level controls included.

Bibliography

Allen, J., Clark, R. and Houde, J.-F. (2014), 'The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry', *American Economic Review* **104**(10), 3365–96.

Arcidiacono, P., Ellickson, P. B., Mela, C. F. and Singleton, J. D. (2020), 'The Competitive Effects of Entry: Evidence from Supercenter Expansion', *American Economic Journal: Applied Economics* **12**(3), 175–206.

Argus Media (2025), 'Argus Oil Market Report', `https://omr.de`.

Armstrong, M. and Chen, Y. (2009), 'Inattentive Consumers and Product Quality', *Journal of the European Economic Association* **7**(2-3), 411–422.

Armstrong, M., Vickers, J. and Zhou, J. (2009), 'Consumer Protection and The Incentive to Become Informed', *Journal of the European Economic Association* **7**(2-3), 399–410.

Asplund, M. and Nocke, V. (2006), 'Firm Turnover in Imperfectly Competitive Markets', *The Review of Economic Studies* **73**(2), 295–327.

Assad, S., Clark, R., Ershov, D. and Xu, L. (2024), 'Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market', *Journal of Political Economy* **132**(3), 723–771.

Atkin, D., Faber, B. and Gonzalez-Navarro, M. (2018), 'Retail Globalization and Household Welfare: Evidence from Mexico', *Journal of Political Economy* **126**(1), 1–73.

Barrett, G. F. and Donald, S. G. (2003), 'Consistent Tests for Stochastic Dominance', *Econometrica* **71**(1), 71–104.

Barron, J. M., Taylor, B. A. and Umbeck, J. R. (2004), 'Number of Sellers, Average Prices, and Price Dispersion', *International Journal of Industrial Organization* **22**(8-9), 1041–1066.

Basker, E. (2005), 'Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail Prices', *Journal of Urban Economics* **58**(2), 203–229.

Basker, E. and Noel, M. (2009), 'The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry', *Journal of Economics & Management Strategy* **18**(4), 977–1009.

Bauner, C. and Wang, E. (2019), 'The Effect of Competition on Pricing and Product Positioning: Evidence From Wholesale Club Entry', *International Journal of Industrial Organization* **67**.

Bento, A. M., Goulder, L. H., Jacobsen, M. R. and von Haefen, R. H. (2009), 'Distributional and Efficiency Impacts of Increased US Gasoline Taxes', *American Economic Review* **99**(3), 667–699.

Bernardo, V. (2018), 'The Effect of Entry Restrictions on Price: Evidence From the Retail Gasoline Market', *Journal of Regulatory Economics* **53**, 75–99.

Berry, S. and Reiss, P. (2007), 'Empirical Models of Entry and Market Structure', *Handbook of Industrial Organization* **3**, 1845–1886.

Borusyak, K., Jaravel, X. and Spiess, J. (2024), 'Revisiting Event Study Designs: Robust and Efficient Estimation', *Review of Economic Studies* **91**(6), 3253–3285.

Bourreau, M., Sun, Y. and Verboven, F. (2021), 'Market Entry, Fighting Brands, and Tacit Collusion: Evidence from the French mobile Telecommunications Market', *American Economic Review* **111**(11), 3459–3499.

Breidenbach, P. and Eilers, L. (2018), 'RWI-GEO-GRID: Socio-Economic Data on Grid Level', *Jahrbücher für Nationalökonomie und Statistik* **238**(6), 609–616.

Bundesamt für Kartographie und Geodäsie (2025*a*), 'Gebietseinheiten 1:5 000 000 (GE5000)', `https://daten.gdz.bkg.bund.de/produkte/sonstige/ge5000/2021/ge5000.utm32s.shape.zip`.

Bundesamt für Kartographie und Geodäsie (2025*b*), 'Verwaltungsgebiete 1:250 000 Stand 01.01. (VG250 01.01.)', `https://daten.gdz.bkg.bund.de/produkte/vg/vg250_ebenen_0101/2021/vg250_01-01.utm32s.shape.ebenen.zip`.

Bundesanstalt für Straßen- und Verkehrswesen (2025), 'Automatische Dauerzählstellen auf Autobahnen und Bundesstraßen ', `https://www.bast.de/DE/Verkehrstechnik/Fachthemen/v2-verkehrszaehlung/zaehl_node.html`.

Bundeskartellamt (2011), 'Sektoruntersuchung Kraftstoffe: Abschlussbericht Mai 2011'.

Busso, M. and Galiani, S. (2019), 'The Causal Effect of Competition on Prices and Quality: Evidence from a Field Experiment', *American Economic Journal: Applied Economics* **11**(1), 33–56.

Byrne, D. P. and De Roos, N. (2019), 'Learning to Coordinate: A Study in Retail Gasoline', *American Economic Review* **109**(2), 591–619.

Byrne, D. P. and Martin, L. A. (2021), 'Consumer Search and Income Inequality', *International Journal of Industrial Organization* **79**, 102716.

Cabral, L., Schober, D. and Woll, O. (2019), 'Search and Equilibrium Prices: Theory and Evidence from Retail Diesel'. Working Paper.

Callaway, B. and Sant'Anna, P. H. (2021), 'Difference-in-differences with Multiple Time Periods', *Journal of Econometrics* **225**(2), 200–230.

Cardoso, L. C., Uchôa, C. F. A., Huamani, W., Just, D. R. and Gomez, R. V. (2022), 'Price Effects of Spatial Competition in Retail Fuel Markets: The Impact of a New Rival Nearby', *Papers in Regional Science* **101**(1), 81–105.

Cengiz, D., Dube, A., Lindner, A. and Zipperer, B. (2019), 'The Effect of Minimum Wages on Low-wage Jobs', *The Quarterly Journal of Economics* **134**(3), 1405–1454.

Chandra, A. and Tappata, M. (2011), 'Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets', *RAND Journal of Economics* **42**(4), 681–704.

Chernozhukov, V., Fernandez-Val, I. and Melly, B. (2013), 'Inference on Counterfactual Distributions', *Econometrica* **81**(6), 2205–2268.

Coglianese, J., Davis, L. W., Kilian, L. and Stock, J. H. (2017), 'Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand', *Journal of Applied Econometrics* **32**, 1–15.

Davidson, R. and Duclos, J.-Y. (2000), 'Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality', *Econometrica* **68**(6), 1435–1464.

Davis, L. W. and Kilian, L. (2011), 'Estimating the Effect of a Gasoline Tax on Carbon Emissions', *Journal of Applied Econometrics* **26**, 1187–1214.

Davis, L. W., McRae, S. and Seira, E. (2022), 'The Competitive Effects of Entry in the Deregulated Mexican Gasoline Market'.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020), 'Two-way Fixed Effects Estimator with Heterogeneous Treatment Effects', *American Economic Review* **110**(9), 2964–2996.

Deutscher Bundestag (2018), 'Bericht über die Ergebnisse der Arbeit der Markttransparenzstelle für Kraftstoffe und die hieraus gewonnenen Erfahrungen', *Drucksache 19/3693* .

Dube, A. (2019), 'Minimum Wages and the Distribution of Family Incomes', *American Economic Journal: Applied Economics* **11**(4), 268–304.

Dubois, P., De Mouzon, O. and Scott-Morton, F. (2015), 'Market Size and Pharmaceutical Innovation', *RAND Journal of Economics* **46**(4), 844–871.

Dubois, P. and Perrone, H. (2017), 'Price Dispersion and Informational Frictions: Evidence from Supermarket Purchases'. Working Paper.

Eckert, A. and West, D. S. (2005), 'Rationalization of Retail Gasoline Station Networks in Canada', *Review of Industrial Organization* **26**(1), 1–25.

ESRI Germany (2025), 'Bundesländer 2017 mit Einwohnerzahl', `https://hub.arcgis.com/datasets/esri-de-content::bundesl%C3%A4nder-2017-mit-einwohnerzahl/explore?showTable=true`.

Fabra, N., Rapson, D., Reguant, M. and Wang, J. (2021), 'Estimating the Elasticity to Real-Time Pricing: Evidence from the Spanish Electricity Market', *AEA Papers and Proceedings* **111**, 425–429.

Federal Ministry for Economic Affairs and Energy (2021), 'Zahlen und Fakten: Energiedaten - Nationale und Internationale Entwicklung'.

Firpo, S., Fortin, N. M. and Lemieux, T. (2009), 'Unconditional Quantile Regressions', *Econometrica* **77**(3), 953–973.

Fischer, K., Martin, S. and Schmidt-Dengler, P. (2024), 'Information and Heterogeneous Tax Pass-through: An Application to Retail Fuel Markets'.

Foros, Ø. and Steen, F. (2013), 'Vertical Control and Price Cycles in Gasoline Retailing', *The Scandinavian Journal of Economics* **115**(3), 640–661.

Gavazza, A. (2011), 'Leasing and Secondary Markets: Theory and Evidence from Commercial Aircraft', *Journal of Political Economy* **119**(2), 325–377.

Goolsbee, A. and 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.

Havnes, T. and Mogstad, M. (2015), 'Is Universal Child Care Leveling the Playing Field?', *Journal of Public Economics* **127**, 100–114.

Hernæs, Ø. M. (2020), 'Distributional Effects of Welfare Reform for Young Adults: An Unconditional Quantile Regression Approach', *Labour Economics* **65**, 101818.

Honka, E., Hortaçsu, A. and Wildenbeest, M. (2019), Empirical Search and Consideration

Sets, *in* 'Handbook of the Economics of Marketing', Vol. 1, Elsevier, pp. 193–257.

Huebener, M., Kuger, S. and Marcus, J. (2017), 'Increased Instruction Hours and the Widening Gap in Student Performance', *Labour Economics* **47**, 15–34.

Ivaldi, M., Jullien, B., Rey, P., Seabright, P. and Tirole, J. (2003), 'The Economics of Tacit Collusion', *European Commission* .

Kilian, L. and Zhou, X. (2024), 'Heterogeneity in the Pass-Through From Oil to Gasoline Prices: A New Instrument for Estimating the Price Elasticity of Gasoline Demand', *Journal of Public Economics* **232**, 105099.

Kim, T. (2018), 'Price Competition and Market Segmentation in Retail Gasoline: New Evidence from South Korea', *Review of Industrial Organization* **53**(3), 507–534.

Kühn, K.-U. and Vives, X. (1995), 'Information Exchanges Among Firms and Their Impact on Competition', *European Communities* .

Lach, S. and Moraga-González, J. L. (2017), 'Asymmetric Price Effects of Competition', *The Journal of Industrial Economics* **65**(4), 767–803.

Lewis, M. (2008), 'Price Dispersion and Competition With Differentiated Sellers', *Journal of Industrial Economics* **56**(3), 654–678.

Li, S., Linn, J. and Muehlegger, E. (2014), 'Gasoline Taxes and Consumer Behavior', *American Economic Journal: Economic Policy* **6**(4), 302–342.

Luco, F. (2019), 'Who Benefits from Information Disclosure? The Case of Retail Gasoline', *American Economic Journal: Microeconomics* **11**(2), 277–305.

Mankiw, N. G. and Whinston, M. D. (1986), 'Free Entry and Social Inefficiency', *RAND Journal of Economics* pp. 48–58.

Martin, S. (2024), 'Market Transparency and Consumer Search - Evidence from the German Retail Gasoline Market', *RAND Journal of Economics* **55**(4).

Maskin, E. and Tirole, J. (1988), 'A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles', *Econometrica* pp. 571–599.

Montag, F., Mamrak, R., Sagimuldina, A. and Schnitzer, M. (2023), 'Imperfect Price Information, Market Power, and Tax Pass-through', *George J. Stigler Center for the Study of the Economy & the State Working Paper* (337).

Montag, F., Sagimuldina, A. and Winter, C. (2023), 'Whom to Inform About Prices? Evidence from the German Fuel Market'.

Moraga-González, J. L., Sándor, Z. and Wildenbeest, M. R. (2017), 'Nonsequential Search Equilibrium with Search Cost Heterogeneity', *International Journal of Industrial Organization* **50**, 392–414.

Nishida, M. and Remer, M. (2018), 'The Determinants and Consequences of Search Cost Heterogeneity: Evidence from Local Gasoline Markets', *Journal of Marketing Research* (3), 305–320.

Noel, M. D. (2007), 'Edgeworth Price Cycles, Cost-Based Pricing, and Sticky Pricing in Retail Gasoline Markets', *Review of Economics and Statistics* **89**(2), 324–334.

Noel, M. D. (2009), 'Do Retail Gasoline Prices Respond Asymmetrically to Cost Shocks?

The Influence of Edgeworth Cycles', *RAND Journal of Economics* **40**(3), 582–595.

Noel, M. D. (2019), 'Calendar Synchronization of Gasoline Price Increases', *Journal of Economics & Management Strategy* **28**(2), 355–370.

Oschmann, S. (2022), 'The Effects of Local Competition on Supermarket Outcomes: Evidence from Superstore Entry using Intra-Market Variation'. Working Paper.

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C. and Yontcheva, B. (2020), 'Information and Price Dispersion: Theory and Evidence', *International Economic Review* **61**(2), 871–899.

Png, I. P. and Reitman, D. (1994), 'Service Time Competition', *RAND Journal of Economics* **25**(4), 619–634.

Png, I. P. and Reitman, D. (1995), 'Why Are Some Products Branded and Others Not?', *The Journal of Law and Economics* **38**(1), 207–224.

Prince, J. T. and Simon, D. H. (2015), 'Do Incumbents Improve Service Quality in Response to Entry? Evidence from Airlines' on-time Performance', *Management Science* **61**(2), 372–390.

Reiffen, D. and Ward, M. R. (2005), 'Generic Drug Industry Dynamics', *Review of Economics and Statistics* **87**(1), 37–49.

Rossi, F. and Chintaguanta, P. K. (2016), 'Price Transparency and Retail Prices: Evidence from Fuel Price Signs in the Italian Highway System', *Journal of Marketing Research* **53**(3), 407–423.

RWI (2023), 'RWI-GEO-GRID: Socio-economic Data on Grid Level - Scientific Use File (Wave 13)', `https://doi.org/10.7807/microm:suf:v3`.

Schaumans, C. and Verboven, F. (2015), 'Entry and Competition in Differentiated Products Markets', *Review of Economics and Statistics* **97**(1), 195–209.

Scope Investor Services (2021), 'Branchenstudie Tankstellenmarkt Deutschland, 2019/2020'.

Seim, K. (2006), 'An Empirical Model of Firm Entry with Endogenous Product-type Choices', *RAND Journal of Economics* **37**(3), 619–640.

Sen, A. and Townley, P. G. (2010), 'Estimating the Impacts of Outlet Rationalization on Retail Prices, Industry Concentration, and Sales: Empirical Evidence from Canadian Gasoline Markets', *Journal of Economics & Management Strategy* **19**(3), 605–633.

Siekmann, M. (2017), 'Characteristics, Causes, and Price Effects: Empirical Evidence of Intraday Edgeworth Cycles', *DICE Discussion Paper No. 252* .

Soetevent, A. R. and Bružikas, T. (2018), 'The Impact of Process Innovation on Prices: Evidence from Automated Fuel Retailing in The Netherlands', *European Economic Review* **110**, 181–196.

Statistical Offices of Germany and the Federal States (2025*a*), '12411-01-01-5: Bevölkerung nach Geschlecht - Stichtag 31.12. - regionale Tiefe: Gemeinden', `https://www.regionalstatistik.de/genesis//online?operation=table&code=12411-01-01-5`.

Statistical Offices of Germany and the Federal States (2025*b*), '13111-08-02-5-B: Sozialversicherungspflichtig Beschäftigte am Arbeits- und Wohnort, Ein- und Auspendler über

Gemeindegrenzen - Stichtag 30.06. - regionale Ebenen', https://www.regionalstatistik.de/genesis//online?operation=table&code=13111-08-02-5-B.

Statistical Offices of Germany and the Federal States (2025*c*), '46251-01-02-4: Kraftfahrzeugbestand nach Kraftfahrzeugarten - Stichtag 01.01. - regionale Tiefe: Kreise und krfr. Städte (bis 01.01.2019)', https://www.regionalstatistik.de/genesis//online?operation=table&code=46251-01-02-4.

Statistical Offices of Germany and the Federal States (2025*d*), '82411-01-03-4: Verfügbares Einkommen der privaten Haushalte einschließlich privater Organisationen ohne Erwerbszweck - Jahressumme - regionale Tiefe: Kreise und krfr. Städte', https://www.regionalstatistik.de/genesis//online?operation=table&code=82411-01-03-4.

Statistical Offices of Germany and the Federal States (2025*e*), 'AI008-1-5: Regionalatlas Deutschland Themenbereich "Erwerbstätigkeit und Arbeitslosigkeit" Indikatoren zu "Arbeitslosenquote, Anteil Arbeitslose" (Gemeindeebene ab 2008)', https://www.regionalstatistik.de/genesis//online?operation=table&code=AI008-1-5.

Statistical Offices of Germany and the Federal States (2025*f*), 'Bruttoinlandsprodukt, Bruttowertschöpfung in den kreisfreien Städten und Landkreisen der Bundesrepublik Deutschland 1992 und 1994 bis 2020', https://statistik.thueringen.de/webshop/pdf/2020/60201_2020_01.xlsx.

Stoline, M. R. and Ury, H. K. (1979), 'Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons Among Means', *Technometrics* **21**(1), 87–93.

Sun, L. and Abraham, S. (2021), 'Estimating Dynamic Treatment Effect in Event studies with Heterogeneous Treatment Effects', *Journal of Econometrics* **225**(2), 175–199.

Tankerkönig (2025), 'Spritpreise in Echtzeit', http://www.tankerkoenig.de/.

Toivanen, O. and Waterson, M. (2005), 'Market Structure and Entry: Where's the Beef?', *RAND Journal of Economics* **36**(3), 680–799.

Varian, H. R. (1980), 'A Model of Sales', *American Economic Review* **70**(4), 651–659.

Whinston, M. D. and Collins, S. C. (1992), 'Entry and Competitive Structure in Deregulated Airline Markets: an Event Study Analysis of People Express', *RAND Journal of Economics* **23**(4), 445–462.

Chapter 3

Indirect Taxation in Consumer Search Markets: The Case of Retail Fuel

Coauthor(s): Simon Martin, Philipp Schmidt-Dengler (both University of Vienna)

Abstract: When consumers have heterogeneous access to price information, they face different observed price distributions and possibly different effective pass-through rates. We estimate a model of consumer search using data from the German retail fuel market. We find that informed consumers face higher effective pass-through rates, with important distributional implications for regulatory and tax policies. Lowering the VAT rate from 19% to 16% decreases transaction prices by 1.9% on average, but disproportionally benefits consumers in high-income markets. A tax-revenue-equivalent excise tax reduction would have benefited consumers more than a VAT cut, thus extending results in public economics to markets with imperfect information.

3.1 Introduction

Consumption taxes are among the most visible components of policy interventions, with an average standard VAT rate of 19.2% in OECD countries (OECD, 2022). They are also a major source of government tax revenue contributing the equivalent of 10% of GDP. They not only transfer resources between consumers, firms, and the government but also redistribute surplus between different groups of consumers and households, e.g., through a progressive income tax schedule. Notably, most of the literature on taxation and tax incidence, dating back to Ramsey (1927) and Mirrlees (1976), operates under the assumption that consumers are perfectly informed about prices.

In light of the recent rise in commodity prices and inflation, policy interventions in the form of tax reductions has drawn substantial attention from the media, academics, and the public. Prices rose particularly in energy markets, which account for a large part of household consumption expenditures.[1] Among the interventions discussed in energy markets, tax cuts in gasoline markets featured prominently and were consequently introduced in several European countries.

In this paper, we study the distributive role of taxation in markets with imperfect consumer information. Specifically, we address the following questions: How are different consumer types affected by tax changes, depending on their access to information and their income? Does the welfare-superiority of ad valorem taxes over unit taxes (Delipalla and Keen, 1992, Anderson et al., 2001*b*) extend to a setting with imperfect consumer information? Considering a model with unit demand, our analysis complements work focusing on the relationship pass-through rates and the curvature of demand (Weyl and Fabinger, 2013, Miravete et al., 2023, Birchall et al., 2024).

The German retail gasoline market provides an ideal setting to study the effect of indirect taxes under imperfect consumer information. First, there is considerable price dispersion, both cross-sectional and intertemporal, despite the fact that the good is physically homogeneous. Second, given this substantial price dispersion, consumer information is key in determining effective prices paid. For example, some consumers use price comparison apps and others do not. Third, in many markets, including the German retail gasoline market, a mix of ad valorem taxes (e.g., a VAT or sales tax) and unit taxes (e.g., an excise tax) is employed, generating substantial tax revenue for the federal government. In 2021, annual excise tax revenues amounted to approximately 33 billion euros (Statistisches Bundesamt, 2023), and VAT revenues from gasoline and diesel reached about 15 billion euros (Deutscher Bundestag, 2022). Finally, we can also exploit variation in the tax scheme. In response to the COVID-19 pandemic, a temporary six-month VAT cut was implemented, namely from 19% to 16%.

Our paper proceeds as follows. We first document stylized facts concerning the German market. We utilize the *value of information* (VOI), which quantifies the difference between the *average* and the *minimum* price at a specific time t in a geographically distinct market m.

[1] In 2022, OECD countries faced inflation rates of 30% for energy while food and prices for other products increased by 7% and 13% respectively (OECD, 2024).

This metric captures both price dispersion and the potential savings for an informed consumer compared to an uninformed one purchasing at a random station. We show that VOI is higher in high-income markets. By linking our price data with car registry data and Google Trends search data, we find that high-income regions are associated with (i) larger cars for which the gains from search are higher and (ii) higher search intensities for gasoline-related keywords on Google. These findings suggest that consumers in different regions might be heterogeneously affected by tax changes due to their access and returns to information.

Indeed, exploiting the aforementioned VAT reduction, we show that high-income regions experience a stronger pass-through of this tax cut. We split the sample of gasoline stations based on the median income per capita of the county in which each station is located and follow a difference-in-differences approach. Prices in above-median income counties decrease *more* relative to below-median income counties after the VAT cut went into effect, suggesting that consumers in above-median income counties benefit about 12% more than those in below-median income counties.

Informed by these reduced-form findings, we quantify different underlying channels through the lens of a structural model with consumer information heterogeneity (Armstrong et al., 2009, Lach and Moraga-González, 2017), allowing for vertical differentiation of gasoline stations (Wildenbeest, 2011). In particular, consumers differ in the number of price quotes they obtain prior to making purchasing decision. This heterogeneity in fixed-sample search arises endogenously in our model due to differences in the costs of obtaining quotes, which vary across consumers and market characteristics such as level of income or competition intensity.[2]

To estimate this model, we propose a two-stage estimation procedure, extending the approach by Wildenbeest (2011). In the first stage, we obtain a non-parametric estimate of the price distribution, conditional on observed market-specific input prices. In the second stage, we match market-level sample moments of the price distribution with those generated by the model to estimate the distribution of search costs (Hong and Shum, 2006, Moraga-González and Wildenbeest, 2008). Intuitively, search costs at the market level are identified indirectly through moments of the observed price distribution. A larger range of observed prices, as well as more drastic adjustments in the price distribution as input prices or market structure change, are indicative of a high fraction of searching consumers, which in turn can be rationalized through low search costs in a given market. We find that estimated search costs are lower in high-income areas as well as markets with denser competition, and decrease over time. All three observations are consistent with the evidence mentioned above.

We then compute counterfactual tax scenarios, motivated by recent tax policy changes in Germany. We perform out-of-sample simulations with a reduced VAT rate of 16% and find that posted prices decrease by 1.92%, corresponding to an average pass-through rate of 77%. Although we estimate our model on pre-pandemic data, this is very much in line with the reduced-form findings from a complementary study by Montag et al. (2023) who compare German to French gasoline stations before and after the VAT cut. They find that prices in

[2]The results in Moraga-González et al. (2017) and Santos et al. (2012) suggest that fixed-sample search captures consumer search behavior well.

Germany fell by 2.06% after the VAT cut.

Our structural model allows us to disentangle this pass-through effect into two channels. First, holding consumer search behavior fixed, the lower tax rate reduces the minimum and average prices while it increases firm profits and price dispersion. Second, consumers respond by intensifying their search as increased price dispersion rewards price comparisons. This allows them to obtain a larger share of the increase to surplus due to lower taxes. Hence, the price falls further. Both channels explain around half of the overall pass-through each. We also show that prices decrease more strongly in high-income markets where consumers search more. In markets in the top decile of the income distribution, the price decrease is 18% stronger than in the bottom decile.

Finally, we investigate how the form of consumption taxation affects outcomes. In particular, we compute the effects of an excise tax reduction such that the total tax revenue equals the revenue obtained under the VAT reduction, i.e. it yields the same outcome from the point of view of the tax authority. We show that, relative to a VAT reduction, prices decrease even stronger when the excise tax is reduced, i.e., when a given tax revenue is financed primarily through VAT.

This result is akin to known results in the public finance literature (e.g., Anderson et al., 2001*b*), although there typically perfect information and elastic demand is assumed. In that case, ad valorem taxes lead to more elastic effective demand. When demand is elastic, estimated pass-through rates crucially depend on the curvature of demand (Weyl and Fabinger, 2013, Miravete et al., 2023, Birchall et al., 2024). In our setting, the consumer preference for cutting the excise tax rather than the ad valorem tax emerges despite imperfect information, equilibrium price dispersion, and no aggregate surplus effect due to inelastic demand. Our result stems from differential distortions in the firm's pricing incentives. Under ad valorem taxes, firms have a stronger incentive to lower prices since the tax authority bears part of the loss in revenue per consumer. In contrast, under excise taxes, the tax revenue per consumer is independent of the price level and hence a price reduction only has a demand effect. In Appendix 3.8, we show that this is a general feature of taxation in homogeneous goods models with consumer search.

Our results have important implications. First, we show how information shapes the heterogeneous effect of tax policy on consumers. We find that differences in search behavior result in considerable heterogeneity in effective pass-through faced. Second, we document that search effort is related to consumers' income. This can inform policymakers about the effective direction and distributional implications of tax changes. Finally, we show that the stated objective of supporting consumers, in particular low-income consumers, reducing the excise tax would have been a more suitable tool than the VAT reduction.

Our paper also contributes methodologically by employing a non-parametric first-stage estimator. Additionally, we demonstrate that characterizing the firm's price distribution (utility) in terms of quantiles allows us to estimate the model based on a large dataset that would otherwise exceed computational capacities, as these quantile expressions directly result in one-dimensional integrals at the market level.

The remainder of the paper is organized as follows. Below we discuss the related literature. Section 3.2 describes the institutional setting and our data. We present descriptive and reduced-form evidence in Section 3.3. Section 3.4 introduces the model and characterizes equilibrium pricing and search behavior. In Section 3.5, we describe our estimation method, and in Section 3.6 the estimation results. In Section 3.7, we conduct and analyze several counterfactual tax experiments, and we conclude in Section 3.8.

Related Literature. Thematically, our paper relates to the vast literature on taxation and tax incidence, going back to Ramsey (1927), see Mirrlees and Adam (2010) for a comprehensive overview.[3] Common themes in this literature include the efficiency of different forms of taxation, as well as overall pass-through rates (Weyl and Fabinger, 2013, Miller et al., 2017, Adachi and Fabinger, 2022, Anderson et al., 2001*a*). Miravete et al. (2023) show the importance of allowing flexibility in a different product demand system, in order to obtain unbiased estimates of pass-through. We contribute to this literature by showing that imperfect price information, modulated through endogenous search, has important consequences concerning pass-through faced by different consumer types in a homogeneous goods market. From an efficiency point of view, the public finance literature has shown that ad valorem taxation is welfare-superior to unit taxes (Delipalla and Keen, 1992, Anderson et al., 2001*b*). We find that also in our setting with imperfect information and unit demand, ad valorem taxes are consumer-surplus optimal despite the lack of output expansion under perfectly inelastic aggregate demand because the revenue-sharing internalization channel of firms is still effective. In contrast to several studies with a macro perspective on pass-through (Bonnet et al., 2024, Gautier et al., 2023, Gelman et al., 2023, Kilian, 2022), we explicitly account for heterogeneity across local markets and consumers. This allows us to quantify several channels arising at the micro level only.

Our paper also contributes to the literature on gasoline markets, surveyed in Eckert (2013) and Noel (2016). Our paper is closely related to Montag et al. (2023) and Genakos and Pagliero (2022) who also investigate pass-through in Germany and Greece, respectively. In a reduced-form manner, Montag et al. (2023) additionally analyze differential effects proxiing consumer information by fuel type. In contrast, we focus on heterogeneous effects across consumer types, depending on their search costs and their relative income. We complement a number of studies on the role of information for consumer rents in gasoline markets (Chandra and Tappata, 2011, Luco, 2019, Martin, 2024, Montag et al., 2021, Nishida and Remer, 2018, Pennerstorfer et al., 2020). Fischer et al. (2024) show that better-informed consumers benefit more from market entry than less-informed consumers. Our structural model allows to quantify this information channel and to discuss the interaction of information and income levels for market outcomes.

Methodologically, we contribute to the literature on estimating search costs (Hortaçsu and Syverson, 2004, Hong and Shum, 2006, Moraga-González and Wildenbeest, 2008, Moraga-González et al., 2013, Wildenbeest, 2011, Honka, 2014, Honka et al., 2019), and, more broadly,

[3]Markets studied empirically include among others liquor (Miravete et al., 2018, 2020) and soda drinks (Dubois et al., 2020).

on markets with imperfect price information and consumer search (Varian, 1980, Burdett and Judd, 1983, Armstrong et al., 2009). Building on the approach introduced by Wildenbeest (2011) and extended by Nishida and Remer (2018), we propose a novel two-stage estimation routine relying on a non-parametric first-step estimator (Li and Racine, 2008). We also show that characterizing the equilibrium price (utility) distribution in terms of quantiles, as in Lach and Moraga-González (2017), has attractive computational features because it allows us to use vectorized Newton's method at the market level. This facilitates the estimation and computation of equilibrium in a model where such calculations would otherwise be computationally infeasible.

3.2 Industry Background and Data

We study heterogeneity in pass-through by different degrees of consumer information in the German gasoline market. Gasoline markets are locally narrowly defined (Bundeskartellamt, 2011, Chandra and Tappata, 2011, Fischer, 2024, Fischer et al., 2024, Martin, 2024, Pennerstorfer et al., 2020). This implies that firms' pricing and pass-through depend on local market structure, demographics, and socio-economic circumstances as well as consumer behavior. The prevalence of many small markets in this industry allows for cross-market comparisons with respect to income levels or consumer information.

We collected data from various sources. First, we use the diesel prices of the universe of German gasoline stations. For our structural model and counterfactual analysis, we utilize data from the pre-COVID and pre-energy crisis period spanning 2015 to 2019. This approach ensures that our results are not distorted by the shocks in 2020. Stations are legally required to report all price changes in real-time to the Market Transparency Unit for Fuel (MTU) of the German competition authority, the Bundeskartellamt. We access this price data through the online portal Tankerkönig (2023).[4] We focus on prices at 5pm on working days when most people fuel (Bundesministerium für Wirtschaft und Energie, 2018). To keep our structural analysis tractable, we restrict our analysis to a 10% random sample of the working days between 2015 and 2019.[5]

The MTU data also include detailed information on the gasoline stations' characteristics. The geographic coordinates of all stations allow us to specify stations' exact locations and to define geographical markets. Information on brand affiliation gives insights into whether stations are vertically integrated into the upstream crude oil and refinery industries (Bundeskartellamt, 2011). The four firms with the highest market share (ARAL, SHELL, TOTAL, ESSO) hold slightly less than 50% of all stations.

Following the literature (Bundeskartellamt, 2011, Fischer, 2024, Fischer et al., 2024, Martin, 2024), we drop all highway stations from the dataset. Even when a highway station is close to a station on a regular road, it is usually considered to belong to a separate market (Bun-

[4] `https://tankerkoenig.de`

[5] We show the robustness of our main results to different times of the day in the Appendix. We also use out-of-sample data for the reduced-form evidence in Section 3.3.3, i.e. data from 2020 and 2021 when the tax changes took place, but do not include this data in our main structural analysis.

deskartellamt, 2011). We follow Fischer et al. (2024) in their procedure to identify highway stations in the data.

Second, we collect data on daily wholesale prices for diesel provided by a private company, Argus Media (2023).[6] Wholesale price data are constructed based on interviews with industry experts and agents who share their wholesale market transaction prices. These wholesale price data already include the excise tax (referred to as *energy tax* in Germany, levied per unit at 47.04 Eurocent per litre, abbreviated as ct/l in the following) but not the VAT of 19%. To understand pass-through in the gasoline industry, we are interested in how changes in wholesale prices map into gasoline retail prices. Comparing the time series of average gasoline prices and the wholesale price data shows they are highly correlated (see Figure 3.E.1 in the Appendix).

We also make use of detailed administrative information on demographic and socio-economic differences across regions in Germany. We obtain data on the income per capita and the share of large cars (cylinder capacity of at least 2000ccm) at the county level (401 “Landkreise”) from Federal Statistical Office (2023).[7] We will exploit the spatial variation in market characteristics to understand heterogeneity in pass-through rates.

Finally, we use data on Google search queries (Google, 2023) on several fueling-related keywords (e.g., diesel, fuel prices, gas station, etc.) at the city level. Aggregating this data to the county level, we later document regional differences in income to differences in search intensity.

We delineate markets using a hierarchical clustering algorithm (Carranza et al., 2015, Lemus and Luco, 2021, Martin, 2024), which generates non-overlapping markets used in our estimation. An advantage of this approach over employing administrative boundaries is that it allows more realistic substitution patterns across artificial boundaries. If instead a fixed radius is drawn around each gasoline station as in Pennerstorfer et al. (2020), market definition does not account for local station density patterns. Moreover, we would not be able to handle the resulting large number of markets in structural estimation and counterfactual equilibrium computation.[8]

The hierarchical clustering algorithm results in 2,328 unique markets including more than 14,000 stations. Table 3.1 presents summary statistics for the key market characteristics. On average, there are around six stations per market, out of which around 40% are classified as “major” stations, and 7% belong to an integrated brand. Figure 3.E.3 in the Appendix shows the distribution of market size. The average maximum distance between a station and the market’s centroid is 4 km. This is in line with market definitions in other papers which use linear or driving distances of one or two miles as market delineations around stations (Chandra and Tappata, 2011, Hastings, 2004, Pennerstorfer et al., 2020).

[6] The same data is also used in, for example, Assad et al. (2024) and Fischer et al. (2024).

[7] `https://regionalstatistik.de`

[8] Figure 3.E.2 in the Appendix displays the market distribution in and around the cities of Aachen and Wuppertal in Germany. The circles’ radii indicate the distance from the market’s centroid, which is the geographical center of a market, to the station farthest away. For our main specification, we parameterize the clustering algorithm with an upper bound of ten stations per market and a maximum distance of ten kilometers between stations.

To match socio-economic variables to the markets, we compute the centroid for each market and assign counties accordingly. As markets are narrowly identified, the vast majority of markets does not include stations from more than one county.
A prominent feature of retail gasoline markets is price dispersion. We calculate three measures of price dispersion, evaluated per market m at a certain time t (5pm on a specific date), given by

$$\begin{aligned} VOI_{m,t} &= E(p_{m,t}) - E_{min}(p_{m,t}) \\ Range_{m,t} &= E_{max}(p_{m,t}) - E_{min}(p_{m,t}) \\ SD_{m,t} &= \sqrt{E(p_{m,t}^2) - E(p_{m,t})^2} \end{aligned}$$

where $VOI_{m,t}$ denotes the *value of information*, i.e., how much a consumer can gain by purchasing at the cheapest (minimum) price $E_{min}(p_{m,t})$ as opposed to the expected price $E(p_{m,t})$ in market m at time t. The price range $Range_{m,t}$ gives the difference between the expected maximum $E_{max}(p_{m,t})$ and minimum price $E_{min}(p_{m,t})$. $SD_{m,t}$ is the market-date-specific standard deviation.
In Table 3.1, we also report summary statistics on prices. Over our sample period, the average price is 116 ct/l. However, there is considerable price dispersion in most markets. On average, consumers can gain 1.6 ct/l when buying at the minimum price instead of the mean price, which is approximately 25% of the margin of a gasoline station in our sample and model. The maximum price in a market is on average 3.4 ct/l higher than the minimum price ($Range_{m,t}$). As approximately 7% of all markets are monopolies, price dispersion in non-monopoly markets is even higher. The degree of price dispersion is slightly larger than, for example, in Fischer et al. (2024) or Pennerstorfer et al. (2020).

Table 3.1: Summary statistics, markets

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
# stations	6.07	2.95	1	10
% Major	0.42	0.27	0	1
% Integrated	0.07	0.14	0	1
% Other	0.51	0.28	0	1
Max(dist)	4.01	2.36	0	10.89
Area	67.91	64.2	0	372.6
Population Density	0.55	0.86	0.04	4.72
GDP per capita	35.7	14.22	15.85	167.21
Mean(price)	116.33	1.85	109.52	131.14
Min(price)	114.71	2.06	108.72	131.14
Max(price)	118.13	2.11	109.52	133.06
S.d.(price)	1.37	0.69	0	7.17
VOI	1.62	1.01	0	6.13
Range	3.43	1.96	0	15.97

Note: This table shows descriptive statistics on characteristics at the market level. In our sample, there are $N = 2,328$ markets, covering 14,000 stations.

3.3 Descriptive Results

3.3.1 Value of Information (VOI)

In this section, we provide first descriptive evidence on how regional differences in socio-economic variables, here measured by income per capita, affect market-level price dispersion. To this avail, we categorize markets into deciles of the income per capita distribution. In Figure 3.1, we show the distribution of price dispersion, measured by the value of information $VOI_{m,t}$, for the markets in the lowest and highest income decile, respectively. Compared to low-income markets, the distribution of price dispersion is shifted to the right in high-income markets. Hence price dispersion tends to be higher in markets with higher income per capita. This could result from consumers searching more intensely in these markets, e.g., because of relatively easier access to price comparison websites or apps.

This pattern holds not only for a cross-section of markets but is also persistent over time. The left panel in Figure 3.2 shows that $VOI_{m,t}$ increases with the wholesale price, and the right panel shows that $VOI_{m,t}$ remains substantially higher for markets in the top decile of the distribution throughout the sample period.

Markets differing in income per capita are likely to also differ in other dimensions such as station density or population density. Hence, we also provide simple linear regressions of market-level price dispersion measures on income per capita and other control variables such as competition proxies (see Table 3.2). They support a significant conditional correlation between price dispersion and income per capita. A 100% increase in income per capita implies an increase in $VOI_{m,t}$ by 0.36 ct/l or more than 20% of the mean respectively. Hence, the gains from being informed are economically relevant higher in high-income markets. We obtain qualitatively similar results for alternative dispersion measures, e.g. the range and standard deviation of market-level prices. The significant relationship between income per capita and the minimum as well as the mean price further indicates that income per capita has an effect on the entire price distribution and not just on very low prices.

Explanations for the heterogeneity in price dispersion for different income levels are multi-fold. Price dispersion can be higher when more consumers search (but also not too many, see Pennerstorfer et al., 2020). Hence, this might be a consequence of different search cost distributions across markets of different income levels. Also, gains from search might be higher in high-income markets as people fuel more (often). We explore some of these possible explanations next.

3.3.2 Larger Cars, Larger Tanks, and Search Intensity

In this section, we establish that higher income regions are associated with (i) larger cars for which gains from search are higher and (ii) stronger search intensity for gasoline-related words on Google's search engine. This explains that indeed search intensity is higher in high-income regions, contributing to the fact that price dispersion there is higher.

First, in the left panel of Figure 3.3, we correlate log income with a county-level ($N = 401$)

Table 3.2: Baseline price regressions, market level

	(1) Mean(price)	(2) Min(price)	(3) S.d.(price)	(4) VOI	(5) Range
Argus Wholesale Price	1.12***	1.07***	0.02***	0.04***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log(Income)	4.05***	3.70***	0.13***	0.36***	0.41***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
# stations	-0.07***	-0.24***	0.05***	0.17***	0.37***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Log(# stations / sqkm)	-0.07***	-0.00	-0.07***	-0.07***	-0.15***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Population Density	-0.28***	-0.31***	0.04***	0.02***	0.06***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	2.31***	7.24***	-1.17***	-4.93***	-5.21***
	(0.07)	(0.07)	(0.03)	(0.04)	(0.06)
Observations	2633692	2633692	2633692	2633692	2633692
R^2	0.900	0.875	0.150	0.177	0.222

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

measure of car size, the share of cars with a cylinder capacity of above 2000ccm.[9] As larger cars consume more fuel, the gains from search are larger in counties with a higher share of such cars. The left panel of Figure 3.3 shows a strong correlation between income and the share of large cars.

Second, the right panel of Figure 3.3 shows that higher income is associated with a higher search intensity for gasoline-related keywords such as Fueling, Gasoline Prices or Gasoline Station on *Google Trends*. Google reports the relative search frequency for keywords, i.e. the share of searches for a keyword instead of the absolute number of searches for a keyword within a region, and standardizes the values to a measure between 0 and 100 to permit a comparison of search intensity across regions or keywords. We construct an index of search, which is the mean search intensity reported for cities within a county across all keywords. The figure shows a significant relation between log income and the standardized search intensity index. We take this as suggestive evidence for more search in high income counties.

3.3.3 Reduced Form Evidence From Tax Changes

In response to the COVID-19 pandemic, Germany implemented a VAT reduction from 19% to 16% (i.e., a reduction by around 16.6%) from July to December 2020. Several lockdowns, disrupted supply chains, and aggregate uncertainty, shocked both the demand supply. Under full pass-through prices would adjust by $\frac{1.16-1.19}{1.19} = -2.52\%$. Montag et al. (2023) analyze this VAT reduction by using France as a control group, and find that average posted diesel prices decrease by 2.06%, which corresponds to a pass-through rate of 82%. Note that since

[9]Approximately 15% of all cars have a cylinder capacity of above 2000ccm. Our results also hold when including the category of cars with 1400ccm to 1999ccm to the group of large cars (64% of all cars have a cylinder capacity of at least 1400ccm).

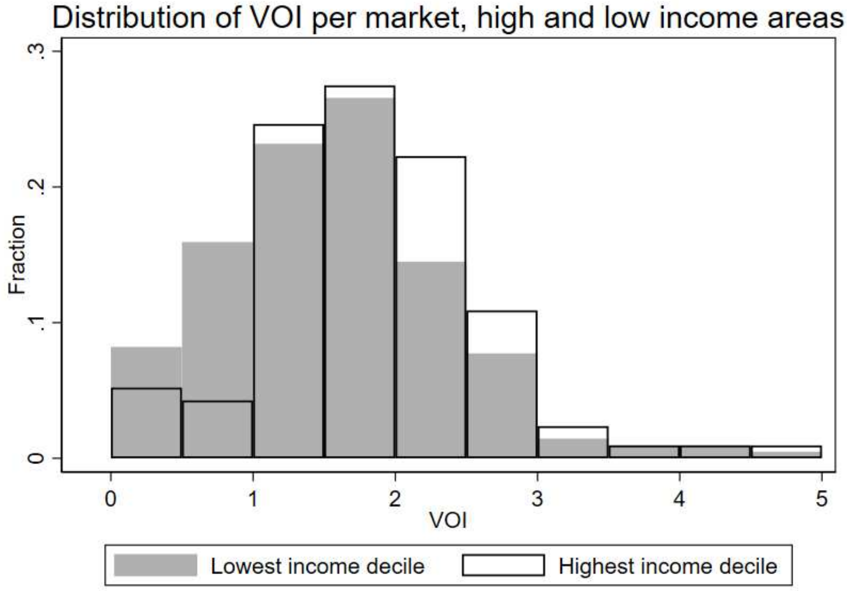


Figure 3.1: Distribution of $VOI_{m,t}$ per market in low and high income p.c. areas

Note: This figure shows the distribution of market-level $VOI_{m,t}$ for the bottom and top decile of the income per capita distribution of markets. $VOI_{m,t}$ is given in ct/l.

the VAT reduction was precisely in response to major changes on the supply and demand side due to the pandemic, it is difficult to isolate the underlying channels.

Instead of only focusing on the average price effect on *all* prices, we are interested in the *heterogeneous* effects of the VAT reduction across markets with different search intensities and income levels. We therefore split the sample of gasoline stations at the median income per capita of the county in which the station is located. We then compare prices of stations in counties with above and below median income and the prices before and after the tax change in a dynamic difference-in-differences estimation.[10]

We estimate the following dynamic regression model:

$$Price_{it} = \alpha_i + \lambda_{st} + \sum_{\tau=-\underline{\tau},\tau\neq-1}^{\overline{\tau}} \mathbb{1}[(Time = \tau)_t] \times \mathbb{1}[Above\ Median_i] + \varepsilon_{it} \qquad (3.1)$$

where $Price_{it}$ is station i's diesel price on date t, α_i and λ_{st} are station and federal state ($N = 16$)-date fixed effects, respectively, and ε_{it} is the error term. The binary variable $\mathbb{1}[Above\ Median_i]$ indicates a station located in an above-median county. We interact this station identifier with weekly bin dummies $1[(Time = \tau)_t]$ to estimate the leads and lags of the treatment effect. We focus on an effect window ($\underline{\tau}$, $\bar{\tau}$) of about ten weeks before/after the change.

This regression setup allows us to identify the effects of tax changes under the parallel trends assumption and the stable unit treatment variable assumption (SUTVA). The former assump-

[10]Note that for this analysis, we also use data outside of our main sample. The post-pandemic time period is omitted from the main analysis for reasons explained in Section 3.2.

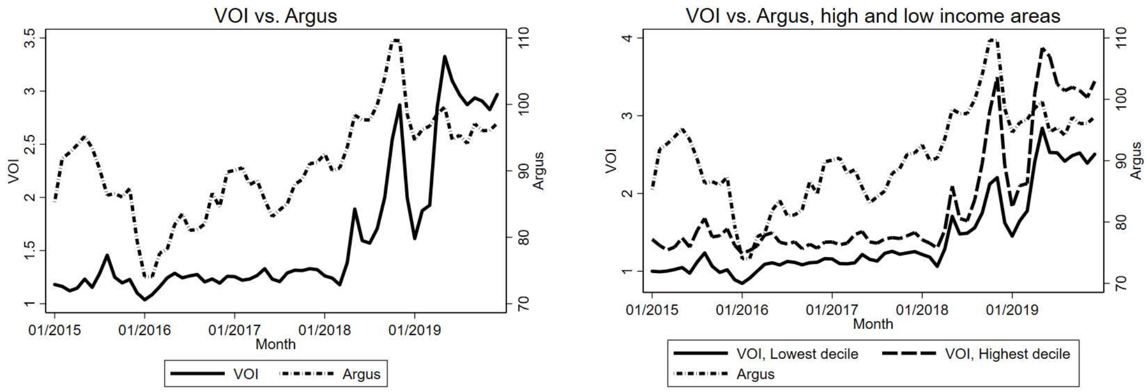


Figure 3.2: $VOI_{m,t}$ and Wholesale Price, including high vs. low income p.c. areas

Note: The left panel shows the time series of Argus wholesale prices and average $VOI_{m,t}$ across markets. The right panel shows the time series of Argus wholesale prices and average $VOI_{m,t}$ for the bottom and top decile of the income per capita distribution.

tion requires that stations in high- and low-income markets would have evolved on similar trends absent treatment. Flat pre-trends serve as suggestive evidence that this assumption is not violated in our setting. The latter assumption implies that there should not be any spillovers in the treatment status across stations. Note that treatment is determined by the local income distribution that remains mostly unaffected in the very short effect windows. Also, stations are unable to self-select, i.e. relocate to markets of different income levels, in response to the policy.

The results are shown in Figure 3.4. Station prices in above-median income counties decrease by 0.25 ct/l relative to below-median income counties. This difference corresponds to about one-tenth of the overall effect to be expected under full pass-through (2.52 ct/l, see Montag et al., 2023). The effect materializes quickly and is persistent over time.

Summarizing our findings so far, we have established that in high-income regions (i) price dispersion is higher, (ii) consumers tend to search more, and (iii) reduced-form evidence suggests that the tax pass-through rate to posted prices is higher.

We will now provide a micro-foundation by estimating a structural model with optimal consumer search. This allows us to disentangle different channels through which tax adjustments operate.

3.4 Model

We consider a setting with vertical differentiation as in Wildenbeest (2011), adjusted for observable input prices and taxes. A finite number of N firms, indexed by i, compete by simultaneously setting prices p_i. There is a continuum of consumers with mass one and unit demand.[11] Firms are vertically differentiated by an observable quality component q_i, which

[11]This assumption on demand is supported by several studies, which find a very low elasticity of demand (Bento et al., 2009, Coglianese et al., 2017, Davis and Kilian, 2011, Levin et al., 2017, Li et al., 2014, Kilian and Zhou, 2024, Knittel and Tanaka, 2021). To provide further support for this assumption, we also show that both traffic and car-related accidents barely respond to the VAT cut in early 2020 (see Figure 3.E.4 and

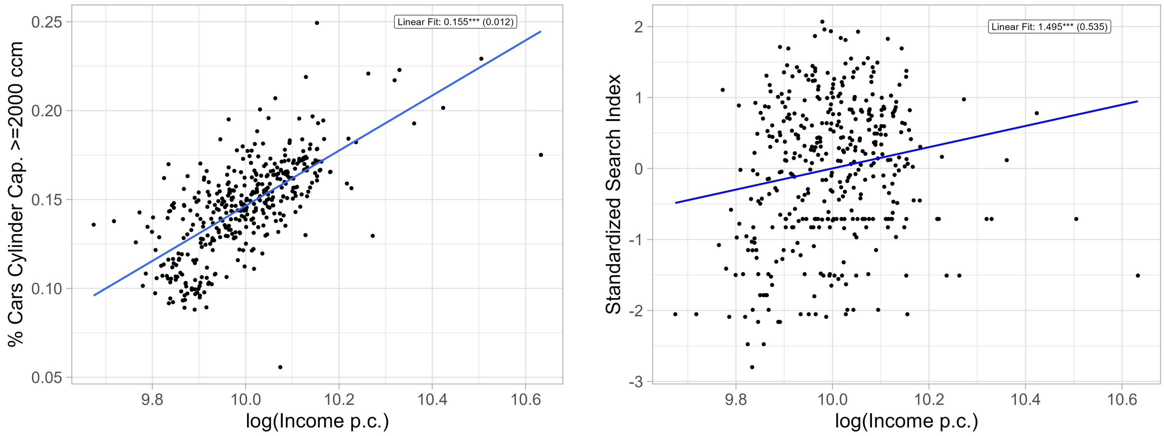


Figure 3.3: Mechanism - Income, Car Type and Search Intensity

Note: This figure correlates log income per capita at the county level with the share of large cars (cylinder capacity $\geq$ 2000 ccm) and a search intensity index based on Google Trends search data. The Google Trends index provides the average search intensity for the following seven words in all cities within a county for which Google Trends reports search intensity data: Tanken (Fueling), Diesel (Diesel), Spritpreise (Fuel Prices), Tankstelle (Gas Station), clever tanken (clever tanken), Benzin (Gasoline), Benzinpreise (Gasoline Prices). If there is not a single city with sufficient search intensity to be reported by Google Trends, we set the search intensity to zero. We also residualize the variable using state fixed effects as Google Trends does not allow for direct comparisons of cities in different states. Finally, we standardize the residualized variable. Linear fits, i.e. the coefficient of an OLS regression of the respective outcome on logged income per capita, are reported in the top right corner. Heteroskedasticity-robust standard errors are reported.

is additively separable from a common quality component q_0, so that the gross utility to consumers is given by $v_i(q_i) = q_0 + q_i$ and the net utility by

$$u_i = v_i(q_i) - p_i = q_0 + q_i - p_i. \tag{3.2}$$

Marginal cost consists of two components: The gross wholesale price for diesel w and the cost of quality provision $r(q_i)$. There is a per-unit tax τ_0, and an ad valorem tax τ_1. The latter is levied on the final product, wholesale fuel, and input costs for quality provision. Taken together, the net revenue per consumer becomes

$$R_i(p_i) = \frac{p_i}{1+\tau_1} - \frac{w + r(q_i)}{1+\tau_1} - \tau_0.$$

Assuming perfectly competitive input markets and constant returns to scale in the production of quality, we have $r(q_i) = q_i$. We rewrite revenue in utility space following Armstrong and Vickers (2001) as

$$\begin{aligned} R_i(p_i) = R_i(u_i) &= \frac{q_0 + q_i - u_i}{1+\tau_1} - \frac{w + q_i}{1+\tau_1} - \tau_0 \\ &= \frac{q_0 - u_i}{1+\tau_1} - c - \tau_0, \end{aligned}$$

where $c = \frac{w}{1+\tau_1}$ equals the net wholesale fuel price. The key insight here is that despite the firms offering asymmetric qualities, we can consider symmetric competition in utility space.

Figure 3.E.5 in the Appendix).

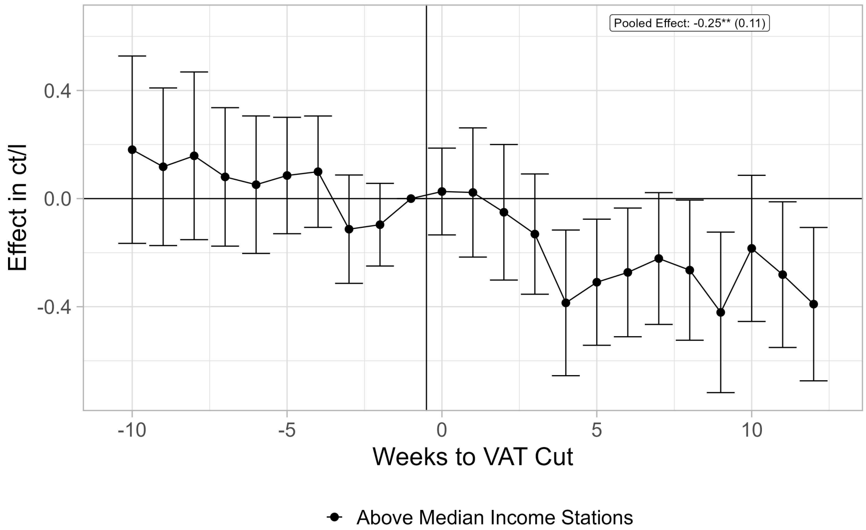


Figure 3.4: Price Effect of VAT Cut on High- Relative to Low-Income Stations
Note: This figure shows the results of a simple difference-in-differences regression of prices on leads and lags of the VAT cut timing interacted with a dummy for stations which are located in counties with an above-median income. We bin leads and lags to weekly bins and use station as well as state-date fixed effects. Standard errors are clustered at the municipality level. 95% confidence intervals are reported. The number in the top-right corner is the simple difference-in-differences estimate (pooled effect).

Consumers are heterogeneous in their information endowment, i.e. the number of prices k (utilities) they observe. A share μ_k observes k prices. The corresponding distribution of information is given by $\{\mu_k\}_{k=1}^N$ and we assume $\mu_1 \in (0,1)$. Before providing a micro-foundation for this heterogeneity below, we characterize optimal firm behavior for a given distribution of consumer information.

As $\mu_1 \in (0,1)$, standard arguments (Varian, 1980, Lach and Moraga-González, 2017) imply that no pure strategy equilibrium exists. Denote the distribution from which utilities are drawn by $L(u)$, which will be symmetric due to the firms' symmetry in utility space. A firm offering utility u_i makes expected profit

$$\pi_i(u_i) = \underbrace{\left(\frac{q_0 - u_i}{1+\tau_1} - c - \tau_0\right)}_{\text{net revenue per consumer}} \underbrace{\sum_{k=1}^{N} \left(\frac{k\mu_k}{N} L(u_i)^{k-1}\right)}_{\text{expected demand}}. \tag{3.3}$$

The equilibrium profit is determined by the profit a firm can make by offering the minimum utility $\underline{u} = 0$, in which case it only sells to consumers who observe just one price. Correspondingly the firm sells quantity $\frac{\mu_1}{N}$:

$$\pi_i^* = \pi^* = \pi_i(0) = \left(\frac{q_0}{1+\tau_1} - c - \tau_0\right) \frac{\mu_1}{N}. \tag{3.4}$$

The equilibrium utility distribution $L(u)$ is then implicitly by the condition that the firm must be indifferent between any utility u and the minimum utility $\underline{u} = 0$, which implies:

$$\pi(u) = \pi^*.$$

Substituting the expressions for $\pi(u)$ and π^* from equations (3.3) and (3.4) yields:

$$\left(\frac{q_0 - u}{1 + \tau_1} - c - \tau_0\right) \sum_{k=1}^{N} \left(\frac{k\mu_k}{N} L(u)^{k-1}\right) = \left(\frac{q_0}{1 + \tau_1} - c - \tau_0\right) \frac{\mu_1}{N}. \tag{3.5}$$

Since (3.5) does not admit a closed-form solution, we rewrite it in terms of quantiles ϕ of L (Lach and Moraga-González, 2017). Let $\xi(\phi) = L^{-1}(\phi) = u$ be the corresponding quantile function. We readily obtain

$$\xi(\phi) = q_0 - \left(\frac{\mu_1 \left(\frac{q_0}{1+\tau_1} - c - \tau_0\right)}{\sum_{k=1}^{N} k\mu_k \phi^{k-1}} + c + \tau_0\right) (1 + \tau_1). \tag{3.6}$$

In order to find the upper bound $\overline{u}$ of the utility distribution, we evaluate (3.6) at $\phi = 1$ and obtain

$$\overline{u} = q_0 - \left(\frac{\mu_1 \left(\frac{q_0}{1+\tau_1} - c - \tau_0\right)}{\sum_{k=1}^{N} k\mu_k} + c + \tau_0\right) (1 + \tau_1)$$

which we can solve for the common quality q_0:

$$q_0 = \bar{u} \frac{\sum_{k=1}^{N} k\mu_k}{\sum_{k=2}^{N} k\mu_k} + (c + \tau_0)(1 + \tau_1).$$

As in Wildenbeest (2011), we obtain the firm-specific price distribution $F_i(p)$ from $u_i = v_i - p_i$, and hence

$$F_i(p) = Pr(p_i \leq p) = Pr(v_i - u_i \leq p) = Pr(u_i \geq v_i - p) = 1 - L(v_i - p).$$

Turning to the consumer side, we define $E_k(u)$ as the expected maximum out of k draws from $L(u)$, i.e. the expected utility of a consumer who samples k firms. The corresponding distribution and density functions are given by:

$$L_k(u) = L(u)^k$$
$$l_k(u) = kL(u)^{k-1} l(u)$$

As $\underline{u} = 0$, we write $E_k(u)$ as

$$\begin{aligned} E_k(u) &= \int_0^{\overline{u}} u l_k(u) du = \int_0^{\overline{u}} u k L(u)^{k-1} l(u) du \\ &= \overline{u} - \int_0^{\overline{u}} L(u)^k du. \end{aligned} \tag{3.7}$$

Consumer heterogeneity in information arises from idiosyncratic search costs in a model of non-sequential search (Burdett and Judd, 1983, Janssen and Moraga-González, 2004, Wildenbeest, 2011). Consumers decide upfront how many prices (utilities) to sample, and subsequently purchase from the firm providing the highest utility in their sample. Following the literature, we assume that the first search is free, but obtaining additional price quotes is costly. Consumers are heterogeneous in their search cost s per price quote, where s is drawn from a continuous and strictly monotone distribution $G(s)$ on $\Re^+$. Consumer choices are optimal given their search cost s and the utility distribution $L(u)$. Thus, a consumer searching k times (weakly) prefers the expected outcome to searching $k' \neq k$ times, i.e.

$$E_k(u) - ks \geq E_{k'}(u) - k's.$$

Since s has full support, this implies a set of cutoff points $\{s_k\}_{k=1}^{N-1}$, each determined by the marginal consumer who prefers k searches to $k+1$ searches:

$$E_k(u) - (k-1)s_k = E_{k+1}(u) - ks_k$$

and hence

$$s_k = E_{k+1}(u) - E_k(u) \tag{3.8}$$

and $s_N = 0$. Therefore, all consumers with $s \in [s_k, s_{k-1}]$ search k times, resulting in shares

$$\mu_k = G(s_{k-1}) - G(s_k), \; k = 2, 3, \ldots, N-1 \tag{3.9}$$

and $\mu_1 = 1 - G(s_1)$ and $\mu_N = G(s_{N-1})$.

The average effective search cost of a type-k consumer is given by

$$E_k(s) = (k-1)\frac{\int_{s_k}^{s_{k-1}} sg(s)ds}{\mu_k}$$

resulting in total average effective search cost for the information distribution $\{\mu_k\}_{k=1}^N$:

$$E_\mu(s) = \sum_{k=1}^N \mu_k E_k(s) = \sum_{k=2}^N (k-1) \int_{s_k}^{s_{k-1}} sg(s)ds. \tag{3.10}$$

3.4.1 Equilibrium

In equilibrium, firms take consumer behavior as given (characterized by their information distribution $\{\mu_k\}_{k=1}^N$), and draw utilities from $L(u;\{\mu_k\}_{k=1}^N)$ in (3.5) (or alternatively, the quantile expression in (3.6)).
Consumers, in turn, take firm behavior as given (characterized by $L(u;\{\mu_k\}_{k=1}^N)$), and search according to the cutoff rule $\{s_k\}_{k=1}^{N-1}$ in (3.8), resulting in $\{\mu_k\}_{k=1}^N$ according to (3.9).
For computing the equilibrium, it is useful to rewrite expressions as in Wildenbeest (2011) to obtain

$$s_k = E_{k+1}(u) - E_k(u) = \int_0^1 u(y)((k+1)y-k)y^{k-1}dy. \tag{3.11}$$

By employing the characterization in terms of quantiles $u(y)=\xi(\phi)$ as in equation (3.6), we can eliminate the dependency on $L(u)$ and write

$$s_k(\{\mu_k\}_{k=1}^N) = \int_0^1 \left[q_0 - \left(\frac{\mu_1\left(\frac{q_0}{1+\tau_1} - c - \tau_0\right)}{\sum_{k=1}^N k\mu_k y^{k-1}} + c + \tau_0\right)(1+\tau_1)\right]((k+1)y-k)y^{k-1}dy$$

and we obtain the equilibrium conditions:

$$\begin{aligned}
\mu_1 &= 1 - G(s_1(\{\mu_k\}_{k=1}^N)) \\
\mu_k &= G(s_{k-1}(\{\mu_k\}_{k=1}^N)) - G(s_k(\{\mu_k\}_{k=1}^N)),\ k=2,3,\dots,N-1 \\
\mu_N &= G(s_{N-1}(\{\mu_k\}_{k=1}^N))
\end{aligned} \tag{3.12}$$

Given that $\mu_N = 1-\sum_{k=1}^{N-1}\mu_k$, this is a $N-1$-dimensional fixed point problem.

3.4.2 Tax Revenue and Welfare

Given a per-unit (excise) tax rate τ_0 and unit demand of a mass 1 of consumers, excise tax revenue is simply

$$TR_0 = \tau_0 \cdot 1$$

and given an ad valorem (VAT) tax rate τ_1, VAT tax revenue is given by

$$TR_1 = \sum_{k=1}^N \mu_k TR_{1,k}$$

where

$$TR_{1,k} = \tau_1 \frac{E_k(p)}{1+\tau_1}$$

and therefore $TR_1 = \frac{\tau_1}{1+\tau_1} E_{trans}(p)$ where the $E_{trans}(p)$ is defined as the weighted expected transaction price:

$$E_{trans}(p) = \sum_{k=1}^{N} \mu_k E_k(p).$$

This results in total tax revenue

$$TR = TR_0 + TR_1 = \tau_0 + \frac{\tau_1}{1+\tau_1} E_{trans}(p).$$

Total welfare is given by

$$W = CS + N\pi^* + TR = q_0 - c$$

so we can obtain expected transaction prices by combining the two preceding equations:

$$E_{trans}(p) = \frac{N\pi^* + \tau_0 + c}{1 - \tau_1/(1+\tau_1)}.$$

3.5 Estimation

Our estimation is based on aggregation to the market-period level observations. We form market-period moments of 'observed' utilities: the mean $E(u_{m,t})$, the standard deviation $sd(u_{m,t}) = E(u_{m,t}^2) - E(u_{m,t})^2$ and the expected maximum $E(u_{max,m,t})$, as well as a fourth moment regarding inter-temporal dispersion. We observe market-level objects upfront and perform all calculations at the market-period level. An overview of our estimation routine is shown in Figure 3.5, and additional details for each step are laid out in the following.

As in Wildenbeest (2011), our starting point is the relationship $u_{it} = v_i - p_{it}$, which can be mapped into the fixed-effects regression

$$p_{it} = \underbrace{\alpha + \delta_i}_{v_i} + \underbrace{\varepsilon_{it}}_{-u_{it}} \tag{3.13}$$

to obtain period-t utility estimates $u_{it} = -\varepsilon_{it}$, which due to the symmetry in utility space can simply be pooled.

We then use a multi-step estimation approach that does not require solving the fairly involved equilibrium fixed-point problem (3.12) at every evaluation of the objective function. Similar to approaches in the literature on auctions and dynamic games, we first estimate (conditional) utility distributions, which can subsequently be used as equilibrium beliefs about firm behavior from the consumers' point of view.

More specifically, we estimate the market-m-specific utility distribution in period t conditional on the wholesale price c, $\hat{L}_{m,t}(u|c)$ non-parametrically, using the method by Li and Racine (2008).[12] We plug this estimated distribution into (3.7) to compute estimated type-k

[12]We use the R package "np" for estimation of the conditional price distributions, see Hayfield and Racine (2008).

Stage 1:

Step 1. Obtain utilities u_{it} through the fixed-effects regression (3.13).

Step 2. Non-parametric estimation of the market-m-specific utility distribution in period t conditional on the wholesale price c, $\hat{L}_{m,t}(u|c)$.

Step 3. Compute estimated type-k specific expected utilities $\hat{E}_{k,m,t}$, and then cut-off points $\hat{s}_{k,m,t}$ in the search costs in (3.8).

Stage 2:

Step 4. For each parameter guess θ, compute the distribution of consumer information types $\{\mu_k\}_{k=1}^n$ by using $G(\hat{s}_{k,m,t};\theta)$ in (3.9), and subsequently quantiles of the utility distribution through (3.6).

Step 5. GMM estimation: Compute the market-period moments in (3.15) and evaluate the criterion function in (3.16). Repeat Steps 4 and 5 until (3.16) is minimized.

Figure 3.5: Estimation routine overview

specific expected utilities $\hat{E}_{k,m,t}$. These estimated expected utilities serve as input in the estimated equilibrium cutoff points $\hat{s}_{k,m,t}$ in the search costs in (3.8). We therefore can treat the cutoff points $\hat{s}_{k,m,t}$ as "data" when estimating the parameters governing search.

We parameterize the search cost distribution as follows, allowing for an annual trend and dependency on market-level observables such as income per capita and the number of stations.[13] Search costs s in market m in year y are assumed to follow a log-normal distribution $s \sim Lognormal(\beta_{m,y}, \sigma_{m,y})$, where

$$\begin{aligned} \beta_{m,y} &= \beta_0 + \beta_1(y - 2014) + \beta_2 \log(inc./cap_m) + \beta_3 \log(n_m) \\ \sigma_{m,y} &= \sigma_0 + \sigma_1(y - 2014) + \sigma_2 \log(inc./cap_m) + \sigma_3 \log(n_m) \end{aligned} \tag{3.14}$$

Thus, we are interested in estimating a parameter vector $\theta = (\{\beta_i, \sigma_i\}_{i=0}^3)$. For each parameter guess θ, we immediately obtain the respective fractions of consumers searching k times using equation (3.9). Then the model-implied objects like $\overline{u}$ and quantiles of the utility distribution are obtained from (3.6). The respective moments are simple one-dimensional integrals at the market-period level, which we can readily compute using the trapezoid method. Computational details are provided in Appendix 3.8.

[13] The semi-parametric approach by Moraga-González et al. (2013) is not directly applicable in our setting due to the additional dependency of search costs on market-level characteristics.

Our moments are given by

$$m(\theta) = \frac{1}{T}\begin{pmatrix} z'[E(\hat{u}_{m,t}) - E(\tilde{u}_{m,t};\theta)] \\ z'[sd(\hat{u}_{m,t}) - sd(\tilde{u}_{m,t};\theta)] \\ z'[\hat{u}_{max,m,t} - E(\tilde{u}_{max,m,t};\theta)] \\ z'\left[\left(E(\hat{u}_{m,t}) - \widehat{E(\hat{u}_{m,t})}\right)^2 - \left(E(\tilde{u}_{m,t};\theta) - \widehat{E(\tilde{u}_{m,t};\theta)}\right)^2\right] \end{pmatrix} \tag{3.15}$$

where $\hat{x}$ denotes the (empirical) mean of x, $\tilde{x}$ denotes the model-implied object x, and the z is an instrument matrix for each of our market-period observations. We use the wholesale price, the number of stations, day-of-the-week dummies, yearly dummies, and market-level demographics as instruments. Our GMM estimator is given by the solution to

$$\underset{\theta}{\operatorname{argmin}}\, m(\theta)'Wm(\theta) \tag{3.16}$$

for a weighting matrix W, e.g., the identity matrix.

Identification. We exploit several sources of variation to identify different elements of the model. A very precise measure of average marginal costs follows right away from the Argus Media input cost data. As in Wildenbeest (2011), the station-specific vertical differentiation component is identified through cross-sectional variation in (average) posted prices. Intuitively, a station that keeps posting higher prices is more likely to offer higher quality (conditional on other observables). The assumption of constant returns to scale in the production of quality does not require us to separately estimate willingness to pay and the cost of quality provision. At the station level, the model allows for vertical differentiation only, thus subsuming also horizontal differentiation elements such as the prominence of the location. A station with higher estimated quality can thus be interpreted as being more conveniently located for consumers *on average*, although not necessarily for all consumers to the same extent. Consumer heterogeneity is reflected solely by the consumer-specific search costs.

Having obtained estimates of quality, we can identify search costs in the following way. Cross-sectional variation in market structure (number of stations and their respective quality components), as well inter-temporal variation in input prices (i.e. the wholesale price) pin down market- and time-period specific gains of searching through the cutoffs in (3.8). Since $\beta_{m,y}$ and $\sigma_{m,y}$ affect both the mean and the variance of the log-normal search cost distribution, the first two out of the four moments in (3.15) would be enough for identifying β_0 and σ_0. The third moment involving $\hat{u}_{max}$ helps in pinning down the location of the distribution. The fourth moment disciplines how the search cost distribution evolves over time, and how strongly search incentives change with changing input prices (and consequently retail prices). Intuitively, this moment ensures that the search habits of consumers remain relatively comparable, even when retail prices increase by over 20% over our sample period.

Finally, the search cost coefficients involving the trend, income per capita, and number of stations, are identified through usage of the respective instruments in the GMM estimation.

3.6 Estimation Results

Our estimation results are shown in Table 3.3.[14] On the consumer side, we estimate a parametric log-normal search cost distribution, resulting from an underlying normal distribution with mean $\beta_{m,y}$ and standard deviation $\sigma_{m,y}$ for each market and year. Search costs are interpreted as the incremental cost of obtaining one additional price quote, including the opportunity costs of time, relative to the costs of filling up an entire tank.

We find that both $\beta_{m,y}$ and $\sigma_{m,y}$ decrease over time, resulting in a 9% decrease in median search costs from 1.10 in 2015 to 1.00 in 2017.[15] Relative to filling up an entire tank of 50l, this implies that the relative costs of obtaining one additional price quote is $1.10 \times 50 \approx 55$ Eurocent, which appears reasonable.

Regarding differences across income groups, we find that mean search costs as well as search cost dispersion are lower in high-income markets. This fits the stylized fact that more search occurs in high-income regions, for example, due to higher opportunity costs of not searching with fuel-intensive cars.

Additionally, station density reduces search costs on average. This can be, for example, due to implicitly lower driving distances between stations in denser markets and a higher probability of sampling gasoline stations on the regular commute.

Table 3.3: Estimation results

Variable	Const.	Trend	log(inc./cap.)	log(# stat.)
Search cost β	1.23 (0.10)	-0.05 (0.00)	-0.10 (0.01)	-0.05 (0.00)
Search cost σ	0.90 (0.09)	-0.01 (0.00)	-0.05 (0.01)	0.01 (0.00)
Med(s), 2015	1.10			
Med(s), 2017	1.00			
Med(s), Low inc./cap.	1.01			
Med(s), High inc./cap.	0.99			
Med(s), Low stat.dens.	1.05			
Med(s), High stat.dens	0.98			
Mean(margin)	6.53			

*Note:*This table shows our baseline estimation results (standard errors in parentheses).

Although we do not match an aggregate margin moment, the estimated margins can serve as a plausibility check of our estimates. With an average of 6.5 ct/l, they are close to those provided in industry reports (Scope Investor Services, 2021) and other papers on the German gasoline market (Assad et al., 2024, Fischer, 2024, Fischer et al., 2024). Relative to average posted prices of 116 ct/l, these tight margins (5.5% of gross retail prices) suggest a competitive market.

[14]Standard errors are obtained from the variance-covariance matrix evaluated at the optimum θ, numerically approximated with finite differences ($h = 10^{-12}$). Thus, the reported standard errors do not consider negligible noise stemming from the first-stage estimation.

[15]Since we are using a log-normal distribution, both β and σ affect the mean and the variance, whereas the median depends on β only. Specifically, $E(s) = \exp(\beta + \sigma^2/2)$ and $med(s) = \exp(\beta)$.

The estimated search cost distributions are primitives of the model. We now discuss how search costs translate into the equilibrium distribution of consumer information, which is a key determinant of firm pricing. The left panel of Figure 3.6, depicts the distribution (across market-date observations) of the mean number of stations k observed per consumer: most consumers observe only one or two prices.

In the next step, we disentangle the average number of observed prices into the different consumer types k at the market level. To illustrate this, the right panel of Figure 3.6 depicts the distribution of consumer types μ_k, for markets with six stations, i.e. the average market size in our sample. On average, around 70% of consumers observe only one station. These consumers purchase at the expected utility $E_1(u_{m,t})$. Due to their relatively high search costs, they still prefer that outcome to searching for cheaper offers (or higher utility). Around 30% of consumers are inclined to compare offers and sample at least two stations, i.e. $k \geq 2$, with most consumers sampling two stations and only very few more than that. The estimated consumer search intensity is comparable to the numbers reported in the survey conducted by the German Federal Ministry of Economics and Technology (Bundesministerium für Wirtschaft und Energie, 2018) and the estimates in Martin (2024).

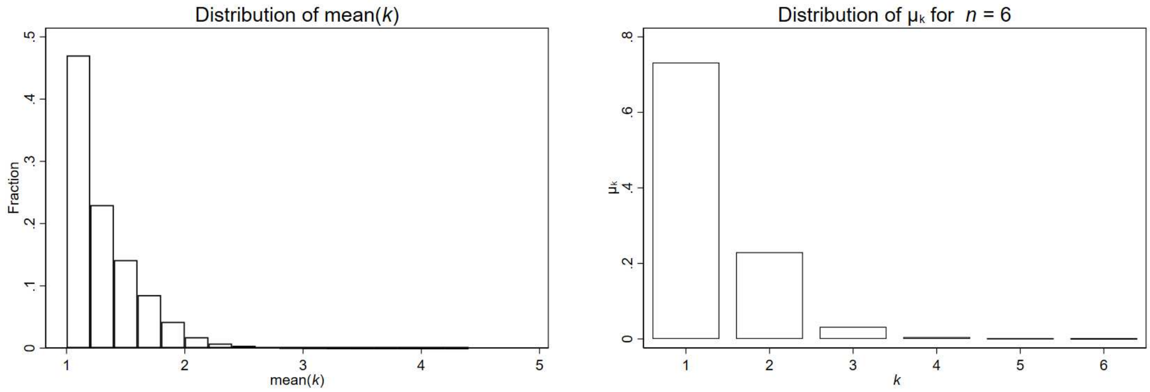


Figure 3.6: Number of stations observed

Note: The left panel shows the distribution of $\bar{\mu}_k$, which is the average number of prices observed in a market. The right panel shows the distribution of the number of prices consumers observe in markets with $N = 6$ firms.

Interpretation of Search Costs. To illustrate the model mechanism, and to interpret the estimated search cost parameters, we compute alternative market outcomes under several counterfactual reductions of search costs. In particular, we quantify the effect of a uniform reduction of 10%, 50%, and 90% of search costs for *all* consumers. Under the maintained assumption of log-normally distributed search costs, a uniform x% reduction in search costs corresponds to adding $log(1-x)$ to β_0 in (3.14). While somewhat ad-hoc, this comparison allows us to highlight the key mechanics of the model.[16]

As mentioned above, the retail gasoline market is already fairly competitive to start with, with average margins in the range of 5-6% of retail prices. The lion's share of retail prices consists of input prices (the wholesale price), and taxes. This implies that rents accrued by

[16]Computational details of our counterfactual analysis are provided in Section 3.7.

firms due to imperfect information are relatively small to begin with.

Table 3.4: Search cost reductions

Spec.	$\Delta E(p)\%$	$\Delta E_{tr.}(p)\%$	$\Delta\Pi\%$	$\Delta\mu_1\%$	$\Delta\mu_2\%$	$\Delta E(k)\%$
Search costs -10%	-0.45	-0.48	-10.71	-9.05	12.70	1.52
Search costs -50%	-2.29	-2.45	-55.03	-49.43	53.61	19.11
Search costs -90%	-2.57	-3.77	-84.84	-79.77	-60.34	132.00

*Note:*This table shows alternative market outcomes (in % relative to the baseline) under several counterfactual search cost reductions, as explained in the main text.

This is confirmed by our counterfactuals in which we reduce search costs, shown in Table 3.4. For each of the three search cost reductions, we report changes (in % relative to the baseline) of average *posted* prices $\Delta E(p)$, average *transaction* prices $\Delta E_{tr.}(p)$, average profits $\Delta\Pi$, the fractions of consumers searching once or twice (μ_1 and μ_2), and the average number of stations sampled per consumer $\Delta E(k)$. For a moderate reduction in search costs (by -10% for all consumers), the resulting effects for consumers are moderate as well. When search costs are lower, consumers are more inclined, on average, to engage in price comparisons, reducing a source of market power. In the new equilibrium, many consumers search twice instead of once. Consumers sample on average 1.52% more stations than before. Firms no longer have monopoly power over those consumers who previously sampled one firm only. Now they compete with a second firm sampled by these consumers, leading to a reduction of posted prices by 0.45%. Since more consumers search, the effect on transaction prices is even slightly stronger. Firm margins were already fairly low to start with. Selling at prices that are 0.48% lower than before reduces firm profits by more than 10%.

A similar pattern emerges for the more drastic reductions in search costs (by 50% or even 90%). Even for the extremely strong reduction of search costs by 90% posted prices decrease by 2.57% only. Compared to the -50% cost reduction, however, *transaction* prices respond considerably. This suggests that firms are competing fiercely under the -50% reduction already. The main gain for consumers then comes mostly from searching for more alternatives, and no longer from putting additional price pressure on firms directly. Relative to the baseline, the average number of sampled firms more than doubles (specifically, increases by 132%). Indeed, both the fraction of consumers searching once or twice drastically decreases, and most consumers search substantially more. In such a competitive market, firms' profits are eroded almost entirely (an 85% decrease relative to the baseline). Nevertheless, the posted prices decreased by 2.57% only. As such, decreasing search costs contributes to shifting rents from firms to consumers.

These considerations suggest that although information frictions are clearly important components of these markets, there is limited scope for drastic changes in prices due to changes in consumer search behavior. Taxes, in contrast, are a major cost driver in this industry, and hence are likely to have stronger effects on prices, as we will elaborate on in the next section.

3.7 Counterfactual analysis

Having obtained estimates of the model primitives, we can now evaluate the distributional implications of public policies, modulated through an endogenous information mechanism. Motivated by recent tax changes that were actually implemented in Germany recently, as described in Section 3.3.3, we compute counterfactual effects of several tax policies. We subsequently show how consumer information influences the impact of these taxes on different groups of consumers. Furthermore, we demonstrate how total alternative revenue-neutral policies have differential effects depending on whether taxes are levied ad valorem or per unit.

We introduce a measure of much *more* consumers in the highest income decile benefit, relative to those in the lowest income decile. Denote by $\Delta E^{inc}(p)$ the relative price change faced by consumers in income decile inc where $inc \in \{low, high\}$. We then define

$$\gamma = 100\frac{\Delta E^{high}(p)}{\Delta E^{low}(p)} - 100.$$

For our counterfactual, we proceed as follows. Based on our estimates of the structural parameters, we compare the status-quo to counterfactual equilibria (see Section 3.4.1). Computing counterfactual equilibria is a relatively involved $N-1$-dimensional fixed-point problem at the market level, solving for the distribution of consumer types $\{\mu_k\}$ in (3.12). This is facilitated by parallelizing at the market-period level.

3.7.1 VAT reduction

Consider the VAT reduction from 19% to 16%, where Montag et al. (2023) find an average price effect of -2.06%.

To illustrate the importance of information frictions and endogenous search behavior, we separately consider the short-term and long-term consequences of the tax policy change. In the short-term, we allow only firms to adjust their prices responding to the tax change while holding the distribution of consumer types $\{\mu_k\}_{k=1}^{N}$ fixed, according to the equilibrium in the baseline specification. As such, this outcome represents only a partial equilibrium analysis, since consumer behavior remains fixed. In the long-term, we allow consumers to adjust their search behavior, such that consumers and firms are again both acting optimally vis-á-vis each other.

In Table 3.5, we depict several outcomes of interest, taking the average across all our markets. Δ_{short} and Δ_{long} denote the relative percentage change over the short (only firms react) and long run (firms and consumers react), respectively. Both in the short and long run, prices decrease and price dispersion increases when the VAT rate is reduced to 16%. Posted prices decrease by 1.92% in the long run. This implies a pass-through rate of 77%. Naturally, the expected minimum price $E_{min}(p)$ and the average transaction price $E_{trans}(p)$ decrease even more, because consumers dis-proportionally purchase at lower prices. Cross-sectional price dispersion measured by the standard deviation $s.d.(p)$ increases, because firms find it relatively more attractive to offer low prices targeted to the informed consumer segment only, since their effective marginal costs are reduced.

In the short run, firms' profit Π increases by 32%, mostly because firms effectively face lower marginal costs and consumer search behavior remains unchanged. This is also highlighted in respective consumer information measure: μ_x and the resulting mean number of stations k observed remains unchanged, by the construction of the short-run counterfactual.

In the long run, however, consumers understand that they should search more in the new environment in which taxes are lower, which leads to lower prices, more price dispersion, and hence higher gains of search. We find that more consumers find it worthwhile to obtain two prices quotes (μ_2 increases by 23%) instead of one price quote only (μ_1 decreases by 16%). This puts additional competitive pressure on the firms since consumers are effectively more price elastic, leading to lower prices than in the short run.

The remarkable differences between short- and long-run effects also highlight the importance of considering information frictions. Not allowing the optimal response of consumers also underestimates the true effects of policy changes.

Table 3.5: Counterfactual results, short and long run

	$\Delta_{short}\%$	$\Delta_{long}\%$
$E(p)$	-0.98	-1.92
$s.d.(p)$	21.06	23.52
VOI	22.28	14.06
$E_{trans}(p)$	-1.10	-2.07
Π	32.79	10.30
μ_1	0.00	-15.93
μ_2	0.00	22.74
mean(k)	0.00	2.05

*Note:*This table shows results for the VAT reduction counterfactual, separately for the short run (where no consumer search adjustment takes place) and the long run (allowing for consumer reoptimization).

We now turn to heterogeneous long-run effects across markets. In Table 3.6, we a breakdown by separately considering only markets in the top (x^{high}) and the lowest decile (x^{low}) in terms of income per capita. Markets with high income per capita experience a stronger price effect, owing to lower search costs, which leads to better-informed consumers.

Although consumers in low-income markets increase their search efforts more (the effect on $mean(k)$ is stronger), high-income areas benefit more from the tax reduction due to the higher baseline levels of searching consumers: The price decrease in high-income areas is around 18% stronger than in low-income areas. Hence our analysis shows that not only average search costs matter for equilibrium outcomes, but the shape of the entire distribution, as also emphasized in Wildenbeest (2011).

The reduced-form estimated effect of 0.25 ct/l (Section 3.3.3) is slightly larger in absolute terms but comparable to our counterfactual results. Note that when the actual tax change was implemented, both supply and demand were disrupted due to the pandemic, possibly leading to further channels not fully picked up by our counterfactual analysis. Nevertheless, the reduced-form estimates and the counterfactual results consistently suggest that high-income areas benefit more from the VAT cut. This is further supported by the evidence shown in Figure 3.7, where we depict the relative price effects for all income deciles, instead of the highest and lowest decile only. Throughout, a consistent pattern of a stronger price effect in high-income areas is evident.

So far, we have shown that consumers benefit from the tax decrease (albeit to a differential degree), and so do firms. We now turn to the government or tax authority. Clearly, total tax revenue (TR) decreases with a VAT decrease. This is shown in the counterfactual results overview in Table 3.7, for the counterfactual with VAT -16% in the short and long run. In the short run, tax revenue decreases by around 4%. In the long run, tax revenues decrease even more (-4.2%), since the tax base, expected transaction prices, is also more strongly affected through the additional consumer search response described above.

Table 3.6: Counterfactual results, low- and high-income areas

	$\Delta^{low}\%$	$\Delta^{high}\%$
$E(p)$	-1.77	-2.09
$s.d.(p)$	25.85	20.38
VOI	16.94	9.85
$E_{trans}(p)$	-1.92	-2.24
Π	13.24	6.77
μ_1	-14.56	-17.49
μ_2	24.73	20.03
mean(k)	2.83	0.59

Note: This table shows results for the VAT reduction counterfactual, separately for low- and high-income areas.

Since we are considering a unit-demand model, prices and taxes are total welfare-neutral transfers between consumers, firms, and the government only. Also, quality production is welfare-neutral under the assumptions above. The only total-welfare relevant quantity is effective search costs $E_{\mu}(s)$, which, from an efficiency point of view, are purely wasteful. In the long run (VAT -16%), consumers search more, leading to an increase of effective search costs $E_{e.}(s)$ by 36.5%. Since search costs are only a negligible fraction of total welfare, total welfare decreases by 0.18%.

3.7.2 Excise Tax Reduction

In the presence of two tax types, as is the case here, either one can be adjusted in order to obtain a certain level of total tax revenue. To put the respective tax changes into perspective, the total reduction in gasoline and diesel tax revenue due to the VAT reduction from 19% to 16% was around one billion euros (Deutscher Bundestag, 2022). Given this substantial foregone revenue from the tax authority's point of view, it is worthwhile to investigate which outcomes could have been achieved by other means. We therefore analyze the potential

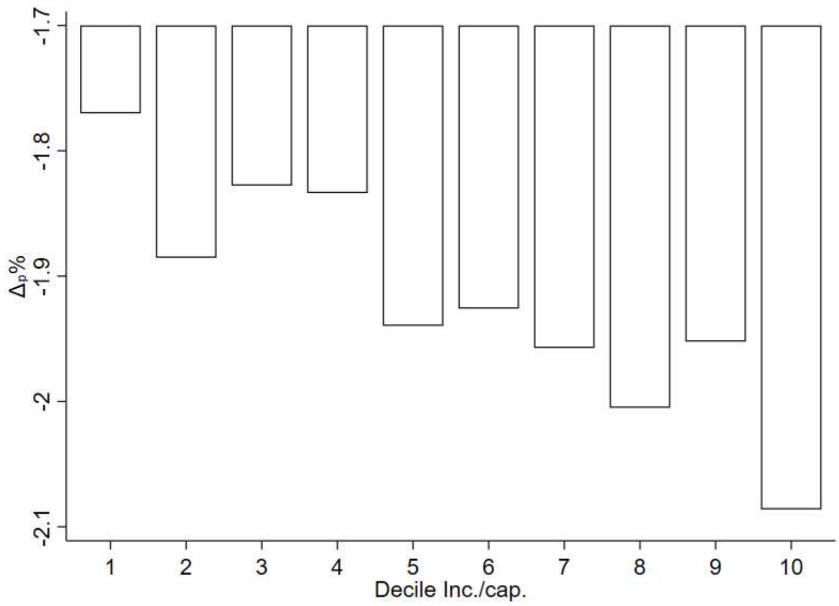


Figure 3.7: VAT change: Long-run price effects per income group

Note: This figure presents the results of the counterfactual analysis when the VAT is lowered from 19% to 16%. The figure shows the percent price change due to the policy for markets from different income deciles.

Table 3.7: Counterfactual results overview

Spec.	$\Delta\ E(p)$	$\Delta\ E_{tr.}(p)$	$\Delta\Pi$	$\Delta\ TR$	$\Delta\ E(k)$	$\Delta\ E_{\mu}(s)$	$\Delta\ W_{inc.}$	γ
VAT -16%, short	-0.98	-1.10	32.80	-4.01	0.00	0.00	0.00	29.43
VAT -16%	-1.92	-2.07	10.30	-4.24	2.04	36.47	-0.18	17.85
Excise -2.4 ct/l	-2.00	-2.16	7.74	-4.23	2.02	35.28	-0.18	17.04
$C0_2$ tax	4.53	5.07	-56.56	12.39	-31.10	-62.05	0.31	24.33

*Note:*This table shows results for several counterfactuals (in % relative to the baseline), as explained in the main text. For the VAT reduction of 16%, both short and long term results are shown.

impact of reducing excise taxes instead of the VAT holding reduction in tax revenue.
An identical tax revenue reduction is achieved when the excise tax is reduced from 47.04 ct/l to 44.68 ct/l, i.e., by 2.36 ct/l or around 5%. The main results are shown in the last row of Table 3.7. Compared to the -16% VAT reduction, the excise tax reduction leads to an even stronger decrease in both posted prices (by 2%) and transaction prices (by 2.2%). Thus, consumers are, on average, even better off under the excise tax reduction than under the VAT reduction. Moreover, the excise tax reduction leads to more equitable outcomes, as is evident in the measure γ - the relative advantage of high-income consumers is only 17% instead of almost 18% under the VAT change. Thus, if the main objective of the tax reduction is promoting consumer welfare, and specifically, welfare of low-income consumers, than reducing excise taxes is the clearly superior tool. These results are in line with the findings of Delipalla and Keen (1992) and Anderson et al. (2001*b*) for markets with perfectly informed consumers. The intuition from these papers carries over in the following way. While firms fully internalize the reduction in revenue from a price decrease under a unit tax regime, the loss in revenue is shared with the government under an ad valorem tax. In Appendix 3.8,

we demonstrate that this effect is not unique to our empirical application but rather, it is a general feature of markets with imperfectly informed consumers.

3.7.3 Other Counterfactuals

In Appendix 3.8, we also demonstrate that all our results remain robust when examining different times of the day. Across various specifications and income deciles, the reduction in VAT is predicted to result in a price decrease ranging between 1.7% and 2.1%. Moreover, the effect size consistently appears to be higher in high-income markets, and a reduction in excise tax would have been preferable from the consumer's perspective.

Our model can also be used to examine other counterfactual policies and their distributional impact across heterogeneously informed consumers and income groups. In Appendix 3.8, we analyze another policy change. Specifically, on January 1st, 2021, the temporary VAT reduction from 19% to 16% expired, coinciding with an increase in the CO_2 price. We investigate this natural experiment using both the reduced-form and structural methods outlined previously in Appendix 3.8. Once again, we find qualitatively similar patterns and results to those observed for the tax counterfactuals explored earlier. For a theoretical and comprehensive treatment of taxation in markets with imperfect consumer information, we refer readers to Appendix 3.8.

3.8 Conclusion

The contribution of regulatory interventions to the efficient allocation of resources is one of *the* central themes in economics, especially in the view of rising commodity prices and inflation. Our study shows an important channel that modulates the effectiveness and distributional consequences of taxation, namely through endogenous information acquisition by consumers. Specifically, we apply a model of non-sequential consumer search to the German retail fuel market, in which cross-sectional price dispersion is a central feature. We find that search costs are decreasing over time. Moreover, search costs are lower in markets with high income per capita than in markets with low income per capita. These results are very well in line with reduced-form evidence.

Endogenous search for prices leads to an atypical form of price discrimination. Although each firm posts one price only and does not discriminate directly, consumers differ in the number of price quotes they obtain (chosen endogenously given their respective search costs). Hence, they also differ in their expected transaction prices. A consumer who samples only one firm observes one price realization only, whereas a consumer who samples ten firms may pick the cheapest out of these ten. This implies that consumers also differ in the *effective* pass-through rates they are faced with. According to our structural estimates, consumers with better access to information pay lower prices, but also their effective pass-through rates are higher.

Based on our model estimates, we compute a counterfactual in which the VAT rate is reduced from 19% to 16%. We find that posted prices decrease by 0.98% in the short run and by

1.92% in the long run, which implies an average pass-through rate of 77%. The long-run effect is stronger due to an adjustment in the endogenous information acquisition by consumers: searching for cheap offers becomes more attractive, putting additional competitive pressure on the firms.

Separately analyzing markets with high and low incomes per capita, respectively, we find that the price reduction following the VAT change is stronger in markets with high per capita income. The main reason is that search costs tend to be lower in these areas. Thus, our analysis shows that the information channel has first-order distributional consequences that should be taken into account by policymakers.

We also show that an excise tax reduction would have been preferable from a consumer welfare point of view. Thus, our findings extend existing results from the public finance literature to a setting with imperfect information, and in which otherwise relevant total demand effects are inactive due to very low aggregate demand elasticity.

As a final note, we mention a limitation that our comparison of different tax types shares the public finance literature, namely considering the tax revenue obtained through different tax types as identical from the tax authority's point of view. This stands in contrast to the legislation in many jurisdictions, according to which for example VAT is part of a different revenue stream than excise taxes levied through gasoline sales. Some of these revenue streams are earmarked for certain expenditures and hence the authority cannot simply transfer tax revenue obtained through different channels, as assumed in our study. We nevertheless believe that our paper is informative about optimal tax design and leave these considerations for future research.

Appendix A: Computational Details

For computation, it is convenient to calculate model-implied objects using integration by parts as follows:

$$
\begin{aligned}
E(\tilde{u};\theta) &= \int_{\underline{u}}^{\overline{u}} ul(u)du = \overline{u} - \int_{\underline{u}}^{\overline{u}} L(u)du \\
E(\tilde{u}^2;\theta) &= \int_{\underline{u}}^{\overline{u}} u^2 l(u)du = \overline{u}^2 - 2\int_{\underline{u}}^{\overline{u}} uL(u)du \\
sd(\tilde{u};\theta) &= \sqrt{E(\tilde{u}^2;\theta) - E(\tilde{u};\theta)^2} \\
E(\tilde{u}_{max};\theta) &= \int_{\underline{u}}^{\overline{u}} ul_{max}(p)du = \overline{u} - \int_{\underline{u}}^{\overline{u}} L(u)^N du
\end{aligned}
$$

with respective sample analogues:

$$
\begin{aligned}
E(\hat{u}_{m,t}) &= \frac{1}{N}\sum_{i=1}^{N_{m,t}} u_{i,m,t} \\
sd(\hat{u}_{m,t}) &= \sqrt{\frac{1}{N}\sum_{i=1}^{N_{m,t}} u_{i,m,t}^2 - E(\hat{u}_{m,t})^2} \\
\hat{u}_{max,m,t} &= \max\left(\{u_{i,m,t}\}_{i=1}^{N_{m,t}}\right)
\end{aligned}
$$

Additionally, we construct a moment capturing inter-temporal and cross-sectional variation based on long-term average objects, i.e.,

$$
\begin{aligned}
\widehat{E(\hat{u}_{m,t})} &= \frac{1}{T}\sum_{i=1}^{T} E(\hat{u}_{m,t}) \\
\widehat{E(\tilde{u};\theta)} &= \frac{1}{T}\sum_{i=1}^{T} E(\tilde{u};\theta)
\end{aligned}
$$

Appendix B: Carbon price (CO_2) Tax

Following up on the reduced form evidence described in Section 3.3.3, the second tax change we observe took place on January 1st 2021, when the VAT rate decrease was undone (i.e., increased back from 16% to the initial 19%), and simultaneously a carbon tax was introduced. The carbon price is 25 Euro per tonne of CO_2, which implies a per-unit tax of 6.69 ct/l for diesel (7.14 including VAT, see Montag et al., 2023). Using again France as control group, Montag et al. (2023) estimate a *joint* pass-through rate of both tax changes of 86% for diesel. Under full pass-through, prices should increase by 9.96% or 10.75 ct/l. We repeat our difference-in-differences estimation for high-and low-income stations using the regression (3.1), where we analyze the income heterogeneity in a reduced-form difference-in-differences setting, comparing above-median income to below-median income stations. The results are shown in an event-study fashion in Figure 3.B.1. Again, we find a significantly stronger and persistent response in high-income areas. Posted prices in above-median income counties increase by 0.41 ct/l relative to below-median income counties. Though, note that this is the joint, reduced-form effect of the simultaneous VAT increase and the CO_2 cut. However, the absolute effect size exceeds the effect of the VAT cut as shown above.

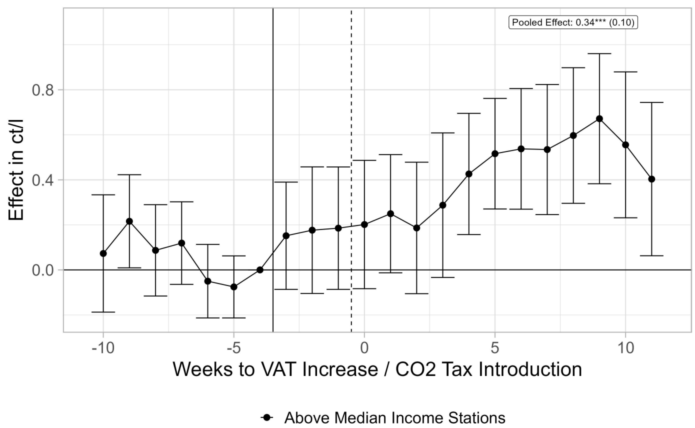


Figure 3.B.1: Price Effect of VAT Increase/CO_2 Tax Introduction on High- Relative to Low-Income Stations

Note: This figure gives the results of a simple difference-in-differences regression of prices on leads and lags of the VAT cut timing interacted with a dummy for stations which are located in counties with an above-median income. We bin leads and lags to weekly bins and use station as well as state-date fixed effects. Standard errors are clustered at the municipality level. 95% confidence intervals are reported. The pooled effect in the top-right corner gives the simple difference-in-differences coefficient where we use the solid vertical line as the effective treatment timing as in Montag et al. (2023) since anticipatory effects were observable.

Analogous to the counterfactual analysis we conducted in Section 3.7, we now simulate a CO_2 tax in a counterfactual manner in our estimated model. In contrast to the VAT, the carbon tax is a non-proportional tax and mathematically is equivalent to an increase of the excise tax. As described above, the carbon price is 6.69 ct/l.

An overview of the counterfactual results is shown in Table 3.7. Posted prices increase by about 4.5%. This seems reasonable in view of the estimates of Montag et al. (2023), who

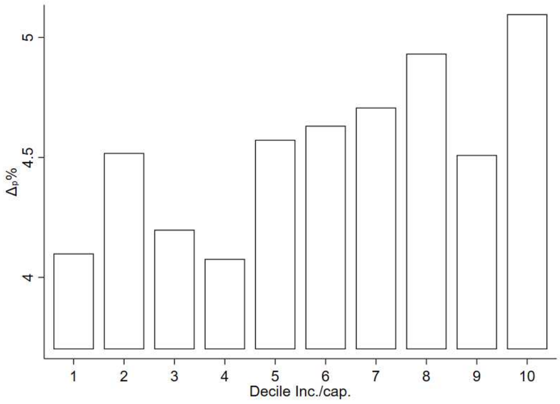


Figure 3.B.2: CO_2 tax change: Long-run price effects per income group

Note: This figure gives the results of the counterfactual analysis when the CO_2 tax of 25 Euro per tonne is introduced. The figure gives the percent price change due to the policy for markets from different income deciles.

investigate a simultaneous VAT increase. As price dispersion decreases under the CO_2 tax, the incentives to search are reduced, leading to fewer searches in equilibrium. This dampens the otherwise even stronger price effect.

Figure 3.B.2 displays the differences in pass-through depending on the market-level income p.c. Again, pass-through increases with income. Pass-through is about 25% stronger in markets from the tenth income decile in comparison to the markets from the lowest income decile.

Appendix C: Robustness Checks

Table 3.C.1: Estimation results (9am)

Variable	Const.	Trend	log(inc./cap.)	log(# stat.)
Search cost β	1.24 (0.09)	-0.04 (0.00)	-0.06 (0.01)	-0.05 (0.00)
Search cost σ	0.90 (0.09)	-0.01 (0.00)	-0.05 (0.01)	0.01 (0.00)
Med(s), 2015	1.65			
Med(s), 2017	1.51			
Med(s), Low inc./cap.	1.52			
Med(s), High inc./cap.	1.49			
Med(s), Low stat.dens.	1.57			
Med(s), High stat.dens	1.48			
Mean(margin)	9.88			

Note: This table shows our baseline estimation results (standard errors in parentheses).

Table 3.C.2: Counterfactual results overview (9am)

Spec.	$\Delta\ E(p)$	$\Delta\ E_{tr.}(p)$	$\Delta\Pi$	$\Delta\ TR$	$\Delta\ E(k)$	$\Delta\ E_{\mu}(s)$	$\Delta\ W_{inc.}$	γ
VAT -16%, short	-0.86	-0.97	22.45	-4.05	0.00	0.00	0.00	22.91
VAT -16%	-1.83	-1.98	7.76	-4.30	-0.50	53.69	-0.28	11.95
Excise -2.4 ct/l	-1.90	-2.06	5.49	-4.19	-0.65	52.30	-0.27	12.40
$C0_2$ tax	5.14	5.79	-22.71	12.55	-29.80	-56.82	0.30	19.92

*Note:*This table shows results for several counterfactuals (in % relative to the baseline), as explained in the main text. For the VAT reduction of 16%, both short and long term results are shown.

Table 3.C.3: Estimation results (noon)

Variable	Const.	Trend	log(inc./cap.)	log(# stat.)
Search cost β	1.23 (0.09)	-0.05 (0.00)	-0.08 (0.01)	-0.05 (0.00)
Search cost σ	0.90 (0.09)	-0.01 (0.00)	-0.05 (0.01)	0.01 (0.00)
Med(s), 2015	1.29			
Med(s), 2017	1.17			
Med(s), Low inc./cap.	1.19			
Med(s), High inc./cap.	1.16			
Med(s), Low stat.dens.	1.23			
Med(s), High stat.dens	1.15			
Mean(margin)	7.63			

Note: This table shows our baseline estimation results (standard errors in parentheses).

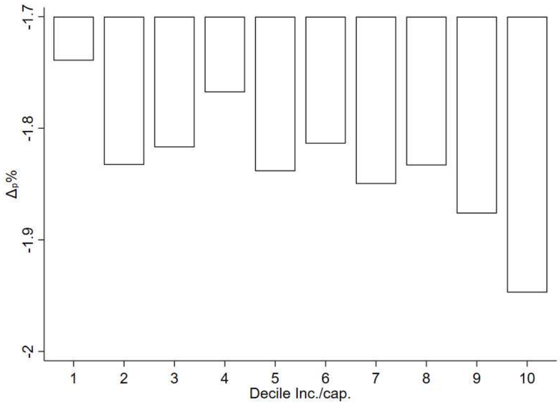


Figure 3.C.1: VAT change: Long-run price effects per income group (9 am)

Note: This figure gives the results of the counterfactual analysis when the VAT is lowered from 19% to 16%. The figure gives the percent price change due to the policy for markets from different income deciles.

Table 3.C.4: Counterfactual results overview (noon)

Spec.	$\Delta\ E(p)$	$\Delta\ E_{tr.}(p)$	$\Delta\Pi$	$\Delta\ TR$	$\Delta\ E(k)$	$\Delta\ E_{\mu}(s)$	$\Delta\ W_{inc.}$	γ
VAT -16%, short	-0.95	-1.07	28.01	-4.02	0.00	0.00	0.00	26.09
VAT -16%	-1.89	-2.05	9.13	-4.25	0.77	38.32	-0.21	15.62
Excise -2.4 ct/l	-1.98	-2.14	6.66	-4.22	0.71	37.02	-0.20	14.23
$C0_2$ tax	4.89	5.49	-38.69	12.50	-31.06	-64.17	0.34	22.63

Note: This table shows results for several counterfactuals (in % relative to the baseline), as explained in the main text. For the VAT reduction of 16%, both short and long term results are shown.

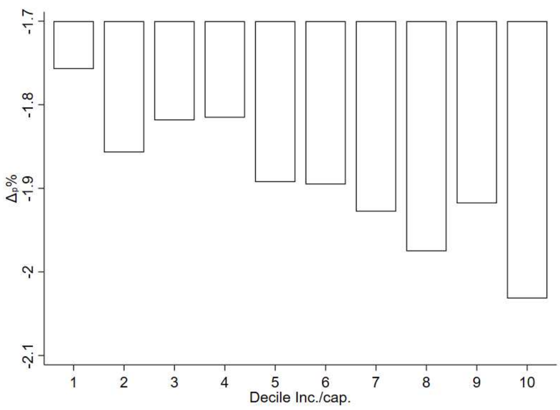


Figure 3.C.2: VAT change: Long-run price effects per income group (noon)

Note: This figure gives the results of the counterfactual analysis when the VAT is lowered from 19% to 16%. The figure gives the percent price change due to the policy for markets from different income deciles.

Appendix D: Taxes in Homogeneous Goods Search Models

For expositional clarity, consider a standard homogeneous goods search model in the spirit of Varian (1980). There are n firms, offering a homogeneous good for which all consumers have willingness to pay $v > 0$. As in Lach and Moraga-González (2017) and our main model, we generalize the distribution of consumer information types, and we assume an exogeneously given distribution $\{\mu\}_{k=1}^{n}$, with the interpretation that μ_k consumers observe k prices. Denote common marginal costs by $c \geq 0$, (per-unit) excise taxes τ_0 and a VAT rate τ_1. For $\mu_1 \in (0, 1)$, a pure strategy equilibrium does not exist, but a mixed strategy equilibrium always exists. In the mixed-strategy equilibrium, firms' profit is determined by selling to loyal consumers only, so

$$\pi = \left(\frac{v}{1+\tau_1} - c - \tau_0\right)\frac{\mu_1}{n}$$

resulting in total-industry profit Π given by

$$\Pi = n\pi = \left(\frac{v}{1+\tau_1} - c - \tau_0\right)\mu_1 l$$

Consumer surplus is simply $CS = v - E_{trans}(p)$, total tax revenue is $TR = \tau_0 + E_{trans}(p)\frac{\tau_1}{1+\tau_1}$ and total welfare is defined through $W = v - c = \Pi + CS + TR$ which we can solve for $E_{trans}(p)$ and obtain

$$\begin{aligned}
v - c &= \Pi + CS + TR \\
v - c &= \Pi + v - E_{trans}(p) + \tau_0 + E_{trans}(p)\frac{\tau_1}{1+\tau_1} \\
E_{trans}(p) &= \frac{\Pi + c + \tau_0}{1 - \frac{\tau_1}{1+\tau_1}} \\
E_{trans}(p) &= \frac{\left(\frac{v}{1+\tau_1} - c - \tau_0\right)\mu_1 + c + \tau_0}{1 - \frac{\tau_1}{1+\tau_1}} = \left(\left(\frac{v}{1+\tau_1} - c - \tau_0\right)\mu_1 + c + \tau_0\right)(1+\tau_1)
\end{aligned}$$

We can use this expression and solve for τ_0 such that total tax revenue TR is constant, i.e., for τ_0 as a function of TR, τ_1, and the other primitives of the model:

$$\begin{aligned}
TR &= \tau_0 + \tau_1\left(\left(\frac{v}{1+\tau_1} - c - \tau_0\right)\mu_1 + c + \tau_0\right) \\
&= \tau_0(1 + \tau_1(1-\mu_1) + \frac{v\mu_1\tau_1}{1+\tau_1} + c\tau_1(1-\mu_1) \\
\tau_0(TR, \tau_1) &= \frac{TR - c\tau_1(1-\mu_1) - \frac{v\mu_1\tau_1}{1+\tau_1}}{1 + \tau_1(1-\mu_1)}
\end{aligned}$$

Similarly, we can then write and simplify $E_{trans}(p; TR, \tau_1)$ as

$$E_{trans}(p; TR, \tau_1) = \frac{v\mu_1 + (1-\mu_1)(1+\tau_1)(c + TR)}{1 + (1-\mu_1)\tau_1}$$

and taking the derivative w.r.t. τ_1, we obtain

$$\frac{\partial E_{trans}(p;TR,\tau_1)}{\partial \tau_1} = -\frac{(v-c-TR)(1-\mu_1)\mu_1}{(1+(1-\mu_1)\tau_1)^2} < 0$$

Thus, holding total tax revenue TR constant, increasing τ_1 (which decreases τ_0 to ensure total tax revenue neutrality) lowers transaction prices. This implies that consumers are better off when a certain level of total tax revenue is financed through high VAT and low excise taxes. Similarly, we can investigate the equilibrium profits π as a function of TR and τ_1

$$\pi(TR,\tau_1) = \frac{v-c-TR}{1+(1-\mu_1)\tau_1}\frac{\mu_1}{n}$$

which is also decreasing in τ_1. The same is true for the lower bound $\underline{p}$ of the price distribution, which also decreases in τ_1.

Appendix E: Additional Figures

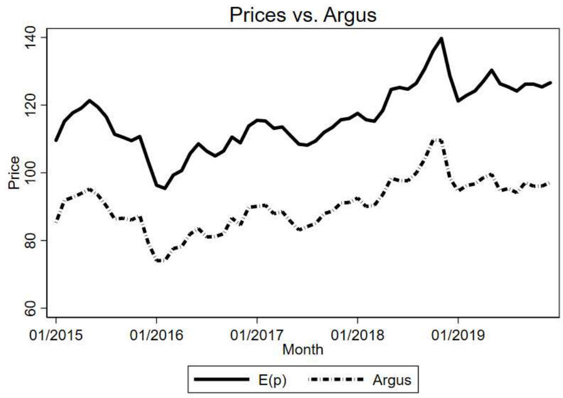


Figure 3.E.1: Retail prices and brent

Note: The figure plots the time series of retail prices across markets and Argus wholesale prices.

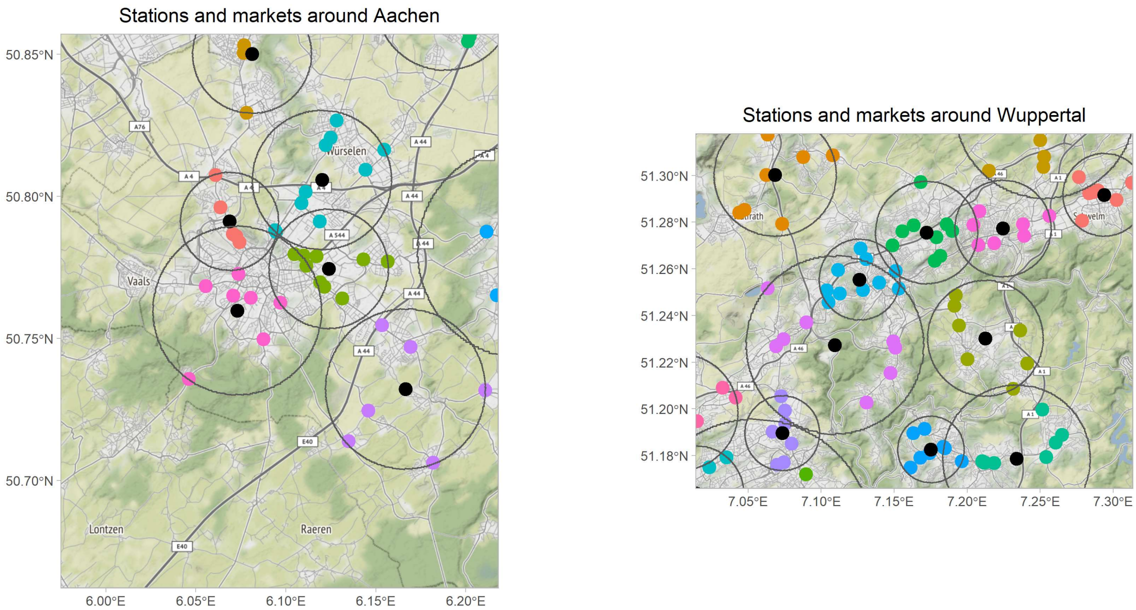


Figure 3.E.2: Illustration of market delineation

Note: The figures represent the market delineation done with a hierarchical clustering algorithm. Different colors represent different markets. Black points represent markets' centroids. Circles' radii have the maximum distance between a market's centroid and a station belonging to the market.

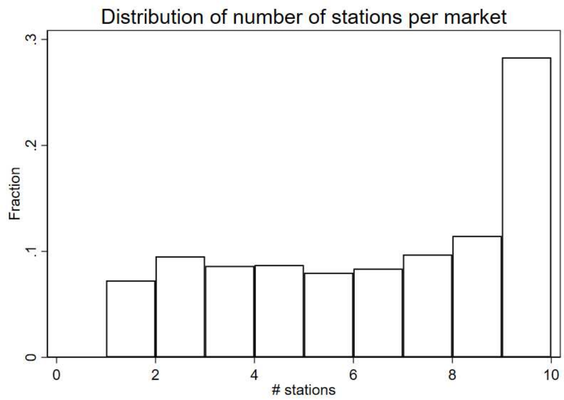


Figure 3.E.3: Distribution of number of stations per market

Note: This figure plots reflects the distribution of market size across markets. Market size is restricted to a maximum number of stations of 10.

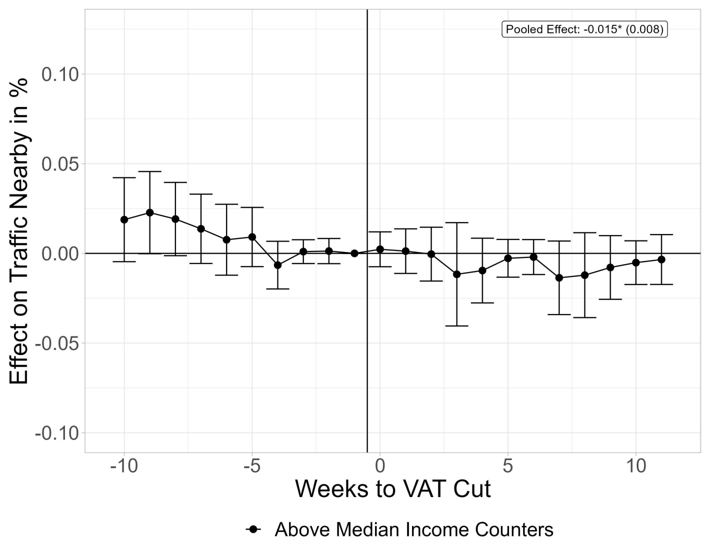


Figure 3.E.4: Traffic Effect of VAT Cut on High- Relative to Low-Income Regions

Note: This figure gives the results of a simple difference-in-differences regression of traffic-counter-level daily traffic on leads and lags of the VAT cut timing interacted with a dummy for counters which are located in counties with an above-median income. We bin leads and lags to weekly bins and use counter as well as state-date fixed effects. There are about 1,500 counters in the sample and only counters which are active over the complete period of the difference-in-differences analysis are included in the estimation. Standard errors are clustered at the municipality level. 95% confidence intervals are reported. The number in the top-right corner is the simple difference-in-differences estimate (pooled effect).

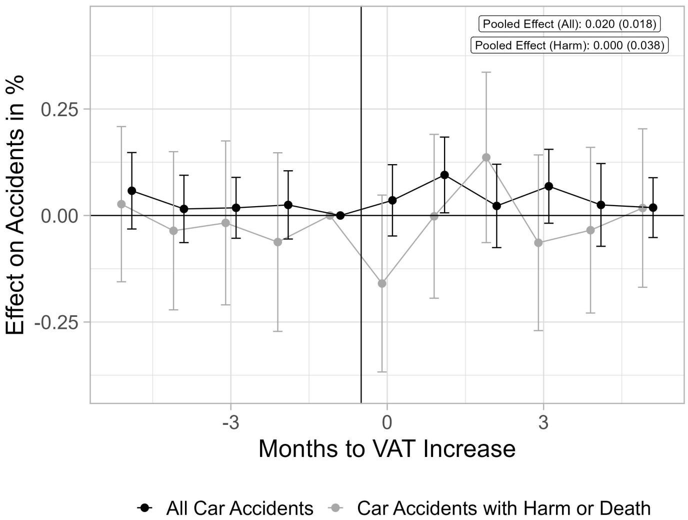


Figure 3.E.5: Accident Effect of VAT Cut on High- Relative to Low-Income Regions

Note: This figure gives the results of a simple difference-in-differences regression of county-level monthly police-reported accidents with personal damage and cars involved on leads and lags of the VAT cut timing interacted with a dummy for counters which are located in counties with an above-median income. We bin leads and lags to weekly bins and use county as well as state-date fixed effects. Standard errors are clustered at the county level. 95% confidence intervals are reported. The number in the top-right corner is the simple difference-in-differences estimate (pooled effect).

Bibliography

Adachi, T. and Fabinger, M. (2022), 'Pass-through, Welfare, and Incidence under Imperfect Competition', *Journal of Public Economics* **211**, 104589.

Anderson, S. P., De Palma, A. and Kreider, B. (2001*a*), 'Tax Incidence in Differentiated Product Oligopoly', *Journal of Public Economics* **81**(2), 173–192.

Anderson, S. P., De Palma, A. and Kreider, B. (2001*b*), 'The Efficiency of Indirect Taxes Under Imperfect Competition', *Journal of Public Economics* **81**(2), 231–251.

Argus Media (2023), 'Argus o.m.r. fuels', https://www.argusmedia.com/en/. Accessed: March 16, 2023.

Armstrong, M. and Vickers, J. (2001), 'Competitive Price Discrimination', *RAND Journal of economics* pp. 579–605.

Armstrong, M., Vickers, J. and Zhou, J. (2009), 'Consumer Protection and the Incentive to Become Informed', *Journal of the European Economic Association* **7**(2-3), 399–410.

Assad, S., Clark, R., Ershov, D. and Xu, L. (2024), 'Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market', *Journal of Political Economy* **132**(3), 723–771.

Bento, A. M., Goulder, L. H., Jacobsen, M. R. and von Haefen, R. H. (2009), 'Distributional and Efficiency Impacts of Increased US Gasoline Taxes', *American Economic Review* **99**(3), 667–699.

Birchall, C., Mohapatra, D. and Verboven, F. (2024), 'Estimating Substitution Patterns and Demand Curvature in Discrete-Choice Models of Product Differentiation', *Review of Economics and Statistics* . forthcoming.

Bonnet, O., Fize, É., Loisel, T. and Wilner, L. (2024), 'Compensation against Fuel Inflation: Temporary Tax Rebates or Transfers?'.

Bundeskartellamt (2011), 'Sektoruntersuchung Kraftstoffe: Abschlussbericht Mai 2011'.

Bundesministerium für Wirtschaft und Energie (2018), 'Evaluierungsbericht zur Markttransparenzstelle für Kraftstoffe'.

Burdett, K. and Judd, K. L. (1983), 'Equilibrium Price Dispersion', *Econometrica* **51**(4), 955–969.

Carranza, J. E., Clark, R. and Houde, J.-F. (2015), 'Price Controls and Market Structure: Evidence from Gasoline Retail Markets', *The Journal of Industrial Economics* **63**(1), 152–198.

Chandra, A. and Tappata, M. (2011), 'Consumer Search and Dynamic Price Dispersion: An Application to Gasoline Markets', *RAND Journal of Economics* **42**(4), 681–704.

Coglianese, J., Davis, L. W., Kilian, L. and Stock, J. H. (2017), 'Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand', *Journal of Applied Econometrics* **32**, 1–15.

Davis, L. W. and Kilian, L. (2011), 'Estimating the Effect of a Gasoline Tax on Carbon Emissions', *Journal of Applied Econometrics* **26**, 1187–1214.

Delipalla, S. and Keen, M. (1992), 'The Comparison Between Ad Valorem and Specific Taxation Under Imperfect Competition', *Journal of Public Economics* **49**(3), 351–367.

Deutscher Bundestag (2022), 'Mindereinnahmen des Bundes bei Aussetzung der Erhebung der Mehrwertsteuer auf bestimmte Energieerzeugnisse'.

Dubois, P., Griffith, R. and O'Connell, M. (2020), 'How Well Targeted Are Soda Taxes?', *American Economic Review* **110**(11), 3661–3704.

Eckert, A. (2013), 'Empirical Studies of Gasoline Retailing: A Guide to the Literature', *Journal of economic surveys* **27**(1), 140–166.

Federal Statistical Office (2023), 'Regionalstatistik', https://regionalstatistik.de. Accessed: March 6, 2023.

Fischer, K. (2024), 'Alcohol Prohibition and Pricing at the Pump', *Journal of Industrial Economics* **72**, 548–597.

Fischer, K., Martin, S. and Schmidt-Dengler, P. (2024), 'The Heterogeneous Effects of Entry on Prices'. Working Paper.

Gautier, E., Marx, M. and Vertier, P. (2023), 'How Do Gasoline Prices Respond to a Cost Shock?', *Journal of Political Economy Macroeconomics* **1**(4), 707–741.

Gelman, M., Gorodnichenko, Y., Kariv, S., Koustas, D., Shapiro, M. D., Silverman, D. and Tadelis, S. (2023), 'The Response of Consumer Spending to Changes in Gasoline Prices', *American Economic Journal: Macroeconomics* **15**(2), 129–160.

Genakos, C. and Pagliero, M. (2022), 'Competition and Pass-Through: Evidence From Isolated Markets', *American Economic Journal: Applied Economics* **14**(4), 35–57.

Google (2023), 'Google Trends', https://trends.google.com/trends/. Accessed: September 29, 2023.

Hastings, J. A. (2004), 'Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California', *American Economic Review* **94**(1), 317–328.

Hayfield, T. and Racine, J. S. (2008), 'Nonparametric Econometrics: The np Package', *Journal of Statistical Software* **27**, 1–32.

Hong, H. and Shum, M. (2006), 'Using Price Distributions to Estimate Search Costs', *RAND Journal of Economics* **37**(2), 257–275.

Honka, E. (2014), 'Quantifying Search and Switching Costs in the US Auto Insurance Industry', *RAND Journal of Economics* **45**(4), 847–884.

Honka, E., Hortaçsu, A. and Wildenbeest, M. (2019), Empirical Search and Consideration Sets, *in* 'Handbook of the Economics of Marketing', Vol. 1, Elsevier, pp. 193–257.

Hortaçsu, A. and Syverson, C. (2004), 'Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds', *The Quarterly Journal of Economics* **119**(2), 403–456.

Janssen, M. C. and Moraga-González, J. L. (2004), 'Strategic Pricing, Consumer Search and the Number of Firms', *The Review of Economic Studies* **71**(4), 1089–1118.

Kilian, L. (2022), 'Understanding the Estimation of Oil Demand and Oil Supply Elasticities', *Energy Economics* **107**, 105844.

Kilian, L. and Zhou, X. (2024), ‘Heterogeneity in the Pass-Through From Oil to Gasoline Prices: A New Instrument for Estimating the Price Elasticity of Gasoline Demand’, *Journal of Public Economics* **232**, 105099.

Knittel, C. R. and Tanaka, S. (2021), ‘Fuel Economy and the Price of Gasoline: Evidence From Fueling-Level Micro Data’, *Journal of Public Economics* **202**, 104496.

Lach, S. and Moraga-González, J. L. (2017), ‘Asymmetric Price Effects of Competition’, *The Journal of Industrial Economics* **65**(4), 767–803.

Lemus, J. and Luco, F. (2021), ‘Price Leadership and Uncertainty about Future Costs’, *The Journal of Industrial Economics* **69**(2), 305–337.

Levin, L., Lewis, M. S. and Wolak, F. A. (2017), ‘High Frequency Evidence on the Demand for Gasoline’, *American Economic Journal: Economic Policy* **9**(3), 314–347.

Li, Q. and Racine, J. S. (2008), ‘Nonparametric Estimation of Conditional CDF and Quantile Functions with Mixed Categorical and Continuous Data’, *Journal of Business & Economic Statistics* **26**(4), 423–434.

Li, S., Linn, J. and Muehlegger, E. (2014), ‘Gasoline Taxes and Consumer Behavior’, *American Economic Journal: Economic Policy* **6**(4), 302–342.

Luco, F. (2019), ‘Who Benefits From Information Disclosure? The Case of Retail Gasoline’, *American Economic Journal: Microeconomics* **11**(2), 277–305.

Martin, S. (2024), ‘Market Transparency and Consumer Search—Evidence from the German Retail Gasoline Market’, *RAND Journal of Economics* **55**(4), 573–602.

Miller, N. H., Osborne, M. and Sheu, G. (2017), ‘Pass-through in a Concentrated Industry: Empirical evidence and Regulatory Implications’, *RAND Journal of Economics* **48**(1), 69–93.

Miravete, E. J., Seim, K. and Thurk, J. (2018), ‘Market Power and the Laffer Curve’, *Econometrica* **86**(5), 1651–1687.

Miravete, E. J., Seim, K. and Thurk, J. (2020), ‘One Markup to Rule Them All: Taxation by Liquor Pricing Regulation’, *American Economic Journal: Microeconomics* **12**(1), 1–41.

Miravete, E. J., Seim, K. and Thurk, J. (2023), *Elasticity and Curvature of Discrete Choice Demand Models*, Centre for Economic Policy Research.

Mirrlees, J. A. (1976), ‘Optimal Tax Theory: A Synthesis’, *Journal of Public Economics* **6**(4), 327–358.

Mirrlees, J. A. and Adam, S. (2010), *Dimensions of Tax Design: The Mirrlees Review*, Oxford University Press.

Montag, F., Mamrak, R., Sagimuldina, A. and Schnitzer, M. (2023), ‘Imperfect Price Information, Market Power, and Tax Pass-Through’.

Montag, F., Sagimuldina, A. and Schnitzer, M. (2021), ‘Does Tax Policy Work When Consumers Have Imperfect Price Information? Theory and Evidence’.

Moraga-González, J. L., Sándor, Z. and Wildenbeest, M. R. (2013), ‘Semi-nonparametric Estimation of Consumer Search Costs’, *Journal of Applied Econometrics* **28**(7), 1205–1223.

Moraga-González, J. L., Sándor, Z. and Wildenbeest, M. R. (2017), ‘Nonsequential Search

Equilibrium with Search Cost Heterogeneity', *International Journal of Industrial Organization* **50**, 392–414.

Moraga-González, J. L. and Wildenbeest, M. R. (2008), 'Maximum Likelihood Estimation of Search Costs', *European Economic Review* **52**(5), 820–848.

Nishida, M. and Remer, M. (2018), 'The Determinants and Consequences of Search Cost Heterogeneity: Evidence from Local Gasoline Markets', *Journal of Marketing Research* **55**(3), 305–320.

Noel, M. D. (2016), Retail Gasoline Markets, *in* 'Handbook on the Economics of Retailing and Distribution', Edward Elgar Publishing, pp. 392–412.

OECD (2022), 'Consumption Tax Trends 2022: VAT/GST and Excise, Core Design Features and Trends', *OECD Publishing* .

OECD (2024), 'Inflation (CPI) (indicator)'.
URL: *https://www.oecd.org/en/data/indicators/inflation-cpi.html*

Pennerstorfer, D., Schmidt-Dengler, P., Schutz, N., Weiss, C. and Yontcheva, B. (2020), 'Information and Price Dispersion: Theory and Evidence', *International Economic Review* **61**(2), 871–899.

Ramsey, F. P. (1927), 'A Contribution to the Theory of Taxation', *The Economic Journal* **37**(145), 47–61.

Santos, B. D. l., Hortaçsu, A. and Wildenbeest, M. R. (2012), 'Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior', *American Economic Review* **102**(6), 2955–2980.

Scope Investor Services (2021), 'Branchenstudie Tankstellenmarkt Deutschland, 2019/2020'.

Statistisches Bundesamt (2023), 'Benzin im Jahr 2022 in Höhe von 12,6 Milliarden Euro versteuert'.
URL: *https://www.destatis.de/DE/Themen/Staat/Steuern/Verbrauchsteuern /energiesteuer.html*

Tankerkönig (2023), https://tankerkoenig.de/. Accessed: September 29, 2023.

Varian, H. R. (1980), 'A Model of Sales', *The American Economic Review* **70**(4), 651–659.

Weyl, E. G. and Fabinger, M. (2013), 'Pass-through as an Economic Tool: Principles of Incidence under Imperfect Competition', *Journal of Political Economy* **121**(3), 528–583.

Wildenbeest, M. R. (2011), 'An Empirical Model of Search with Vertically Differentiated Products', *RAND Journal of Economics* **42**(4), 729–757.

Chapter 4

Industrial Policy in Declining Industries: Evidence from German Coal Mines

Coauthor(s): none

Abstract: Industrial policy is on the rise. However, empirical evidence of how industrial policy shapes technological progress and productivity remains scarce. This paper examines a policy that aimed at boosting industry-wide productivity by subsidizing plant closures in the declining German coal mining industry. Based on newly digitized, mine-level production data, my findings indicate that the policy increased long-run productivity in three distinct ways: First, it facilitated the exit of low-productivity mines. Second, it triggered reallocation towards large, productive mines, especially in firms where the subsidy alleviated financial constraints. Third, firms invested parts of the policy-induced subsidies into machinery and infrastructure of surviving mines. The resulting within-mine productivity gains extended mines' lifespan by six years. In total, the associated reduction in marginal cost exceeded the government subsidies.

4.1 Introduction

Many Western economies have been facing significant changes in their industry composition, resulting from forces such as rising international trade or the advent of new technologies. Industries that were once highly relevant, such as steel production or car manufacturing, have declined for years (e.g., Bekaert et al., 2021, Dechezleprêtre et al., 2023). In a laissez-faire scenario, these industries might simply disappear, but for multiple reasons politicians strive to sustain such industries, or at least steer and decelerate their decline. For example, industrial decline and layoffs are typically regionally concentrated and induces (socio-)economic disparities in space (e.g., Berbee et al., 2024, Gagliardi et al., 2023) as well as political polarization (e.g., Autor et al., 2019, Dippel et al., 2022). Further, keeping an industry alive can serve strategic economic goals and ensure geopolitical independence or can help to overcome transitory causes for the decline.

For these reasons, politicians have an interest in supporting certain industries and pursuing industrial policies. Hence, it is crucial to understand which policies can be implemented in this context to meet the policymaker's goals in the most efficient way. However, while a large share of industrial policies is devoted to promoting declining industries, empirical evidence on the effects of such interventions remains scarce. This is due to the just recent resurgence of industrial policy (Barwick et al., 2024*b*, Juhász et al., 2022, 2023) and the focus of research on industrial policy in growing or new infant industries (Barwick et al., 2024*a*, Harris et al., 2015, Juhász, 2018, Lane, 2022, Manelici and Pantea, 2021, Rodrik, 2004) or placed-based policies after industries have declined (Cingano et al., 2022, Criscuolo et al., 2019).

In this paper, I shed light on the question of how industrial policy can enhance industry-wide productivity in declining industries, thereby actively steering the industry's decline. I study a historical episode of a specific industrial policy in the shrinking German coal mining industry in the 1960s and 1970s. At the time of the policy's introduction, the industry accounted for 4.5% of the national GDP (Federal Statistical Office of Germany, 1965), faced severe import competition with oil, and was set to decline considerably. However, rather than commonly subsidizing the industry's production to decelerate the decline, the government pursued the unconventional strategy of offering closure payments.[1] Through this program, firms closed 25% of the industry's capacity.

I show that this policy led to considerable productivity gains, by triggering the exit of unproductive mines, within-mine productivity growth, and within-firm reallocation towards more productive mines. This episode may have ramifications for the design of current policies in industries that are declining or hold excess capacities. My findings emphasize that the long-term survival of the industry might be achieved by consolidating industry capacities in more productive plants through targeted policies, rather than maintaining all production capacities in all firms via subsidies.

I study the policy's impact using detailed production data in physical units at the estab-

[1] Closure payments have only been used to reduce overcapacities in few other industries, e.g., the EU crop industry (Commission of European Union, 1988), EU fishing industry (Council of European Union, 2006), French steel industry (Raggi et al., 2015), or EU milk industry (Commission of European Union, 2016).

lishment level for the universe of German coal mines, which I newly self-digitized from various archive sources. Employing both reduced-form methods and the structural production approach in the spirit of Ackerberg et al. (2015), De Loecker and Warzynski (2012), and De Loecker and Scott (2022), I demonstrate that the consolidation policy led to a significant productivity increase. Relative to Belgian mines that quarried from the same cross-border coal field and had similar development trajectories before the policy, German mines saw an approximately 10% increase in labour productivity over a ten-year time span after the policy on average.

This productivity rise can be attributed to three almost equally important channels. First, the closure subsidy led to positive selection, i.e., the exit of inefficient mines. I observe a negative effect of a mine's labour productivity and total factor productivity (TFP) on its likelihood of exiting under the policy. Since exit depends on many unobserved factors that possibly correlate with productivity, I use an instrumental variable approach that leverages differences in mines' geology as an exogenous shifter of productivity for the identification of a causal effect. My preferred specification suggests that exit increased labour productivity (TFP) by 3.1% (1.5%).

Second, I show that firms used the closure subsidies to improve the productivity of their remaining mines. I compare the remaining mines of policy uptakers to non-uptakers, that both developed on similar pre-policy outcome trends, in a difference-in-differences design. I find that the subsidies alleviated the financial constraints of treated firms. The policy reduced uptaking coal firms' debt ratios by up to 15 p.p., while simultaneously boosting their stock market values by on average 30% and increasing dividend payouts by on average 20%. As a result, the subsidies induced more investments, which resulted in better infrastructure and technology adoption. Formerly more financially constrained firms responded more strongly to the policy. Overall, these adjustments contributed another 3.3% (4.1%) to industry-wide labour productivity (TFP) gains on average. I further show that estimated marginal costs of mines owned by treated firms fell by around 1.5%, resulting in cost savings that exceeded the government expenditures through the policy. The investments also extended the lifespan of treated mines by six years on average. Workers in treated mines profited through wage increases relative to those in mines of non-uptakers.

Third, by studying heterogeneity in mine characteristics, I show that the policy facilitated within-firm reallocation towards larger, more productive mines in multi-mine firms. This reallocation led to the emergence of a few very large and highly productive mines. Using distribution regressions in the spirit of Chernozhukov et al. (2013), I elicit the full counterfactual mine size distribution absent the policy. I find that absent the policy the largest mine would have been smaller than one-fourth of the treated mines post-policy. The observed concentration of productivity growth in a few large mines aligns with recent evidence which emphasizes that industry-wide productivity gains are often driven by a few rapidly growing 'superstar' firms (Autor et al., 2020, De Loecker et al., 2020, De Loecker and Eeckhout, 2018, Stiebale et al., 2024). This reallocation channel contributed another 3.2% (1.7%) to the industry's labour productivity (TFP) gains.

Whereas the policy increased average industry-wide productivity, it also had significant distributional effects between firms. Recall that the policy's goal was to reduce capacity. In contrast to single-mine firms, multi-mine firms got around this policy target. They earned a premium for mine closures but shifted the full production volume of the closed mines to the remaining mines post-policy. I show that this caused an increasing productivity dispersion in the industry with deteriorating mines at the left tail and improving mines at the right tail of the productivity distribution. Further, I find that policy-uptaking firms revealed higher stock values, dividends, and lower debt ratios after the policy relative to non-uptakers.

The policy also caused a reallocation of output towards mines with cokeries. Cokeries refine coal to coke, a critical input for steel production. Steel production had been a reliable source of demand for coal, consuming around 40% of produced coal (Gatzka, 1996). As coke cannot be substituted with oil, the steel industry did not reduce its demand for coal and coke. Hence, the policy led to a shift in mines' business model along the value chain towards more stable customer markets. Thus, reallocation led mines to adapt to and insure themselves against the upcoming decline in household demand for coal. As a side effect of this reallocation, I show that the policy increased employment in vertically integrated cokeries owned by coal firms.

While the policy laid off or retired 29,000 workers in the short-run (accounting for employment spillovers), it ultimately induced a higher survival rate of mines. A careful back-of-the-envelope calculation suggests that the extended longevity of these mines saved about 20,000 jobs per year over the post-policy horizon. I neither find positive nor negative employment spillovers to other industries in counties where mine closures took place, i.e., mine closures do not cause a deindustrialization outside of the narrowly defined coal industry.

I also thoroughly illustrate that the policy was cheaper than common alternative interventions. First, I show that price subsidies of the same volume as the implemented closure subsidies would have only sustained demand for excess coal production for two years. Moreover, wage subsidies or increased government consumption of excess coal would have quickly been more expensive policies than closure subsidies. I also demonstrate that promoting within-firm mine mergers would not have achieved similar productivity and efficiency gains as the closure subsidy.

My results are informative to policymakers about how to conduct industrial policy in declining industries, in industries with (temporary) costly overcapacities (e.g., milk production), or in which the decline of aggregate output is a policy goal (e.g., non-renewables).

Related Literature. This paper relates to several strands of the literature that motivate research on industrial policy in declining industries.

First, a large literature has documented the misallocation of production across countries (Hsieh and Klenow, 2009, 2014, 2018, Hsieh et al., 2019, Restuccia and Rogerson, 2017), industries (Adamopoulos et al., 2022, Hsieh and Klenow, 2009, 2014, 2018) and firms (Asker et al., 2019, 2020) as an important source of productivity losses. These papers examine the drivers or obstacles of reallocation such as trade liberalization and import/export competition (Pavcnik, 2002), innovation (Hsieh and Klenow, 2009, 2018) or demand fluctuations

(Collard-Wexler, 2013, Allcott et al., 2016). In my paper, I take these insights to declining industries and show that reallocation increases industry-wide productivity. A novel aspect of my findings is that this reallocation occurs within firms across their establishments. My results also support the notion that financial constraints are a key hurdle to reductions in misallocation (Midrigan and Xu, 2014).

Second, I address an extensive literature discussing the pros and cons of industrial policy (Juhász et al., 2022, Juhász and Steinwender, 2023, Rodrik, 2004, 2009). Common concerns about industrial policy include its high cost, lack of precise targeting, and potential to discourage innovation. In my paper, I show that industrial policy can be less costly than expected. I further show that firms adopt their business model along the value chain and adopt new technologies. By looking at exit subsidies in a declining industry, I also provide a new perspective on conducting industrial policy. Up to now, most evidence on the effects of industrial policy has been provided for rising infant industries (Juhász, 2018, Lane, 2022), trade policy (Brandt et al., 2017, De Loecker, 2011, Orr and Tabari, 2024, Pavcnik, 2002, Topalova and Khandelwal, 2011) and standard policies such as price, wage, investment or place-based subsidies (Becker et al., 2010, Criscuolo et al., 2019, Ehrlich and Seidel, 2018, Garin and Rothbaum, 2024, Heblich et al., 2022, Kline and Moretti, 2014, LaPoint and Sakabe, 2021, Siegloch et al., 2024).[2] I demonstrate that closure subsidies can preserve more jobs in the long term compared to price or wage subsidies, assuming a fixed policy budget.

Third, my findings add to a growing literature that investigates the determinants of firm- and establishment-level productivity and productivity dispersion. Various factors have been identified, such as competition intensity (Backus, 2020, Stiebale and Szücs, 2022, Syverson, 2004), ownership (Braguinsky et al., 2015), import competition (Amiti et al., 2019, De Loecker and Warzynski, 2012, Lileeva and Trefler, 2010, Topalova and Khandelwal, 2011), exporting (De Loecker, 2013), manager ability (Rubens, 2023*a*) or FDI (Arnold and Javorcik, 2009, Lu and Yu, 2015). My paper sheds light on the role of policy interventions and financial constraints in declining industries. I find that the policy increased overall productivity as well as productivity dispersion. This allows me to quantify the role of government policies in encouraging productivity growth, which up to now has been less extensively studied (Syverson, 2011), by using the context of German coal mining.

Fourth, I engage with the literature on the impact of consolidation on productivity (Grieco et al., 2018, Rubens, 2023*b*), profitability (Braguinsky et al., 2015), and input and output market power (Guanziroli, 2022, Miller and Weinberg, 2017, Rubens, 2023*b*, Prager and Schmitt, 2021, Schmitt, 2017). By highlighting within-firm reallocation across plants, I provide evidence for a new mechanism through which consolidation policies affect plant- and firm-level productivity as well as markups and wages. In line with Aghion et al. (2015) who show that industrial policy can lead to strong productivity growth, especially in competitive industries, I show that industrial policy in coal mining creates substantial productivity gains in the light of import competition from oil.

[2] An exception is Heim et al. (2017) who look at EU state aid for firms in business crises. They find that such rescue policies improved the survival rate of treated firms.

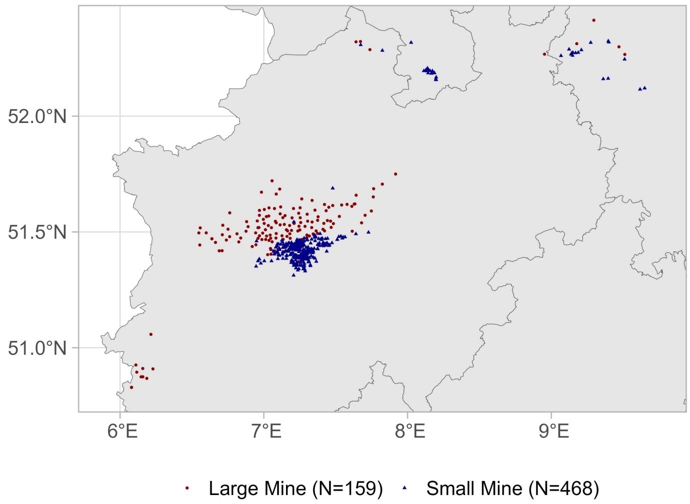


Figure 4.1: Geographical Distribution of Mines

Note: This plot shows the location of all mines which operated after 1947. Large mines are all mines that reported detailed production data. Small mines are all others. Merged mines are shown in their original pre-merger separation. Federal state borders of Northrhine-Westphalia included.

The remainder of the paper is structured as follows: I first explain the institutional setting in Section 4.2. I then describe the data, the TFP and markup estimation, and the empirical analysis in Sections 4.3 and 4.4. In Sections 4.5 and 4.6, I present the results of the paper and discuss potential mechanisms before I provide robustness checks and conclude by discussing the implications of my results in Sections 4.7 and 4.8.

4.2 Institutional Setting

In this section, I give an overview of the German coal mining industry and its decline.

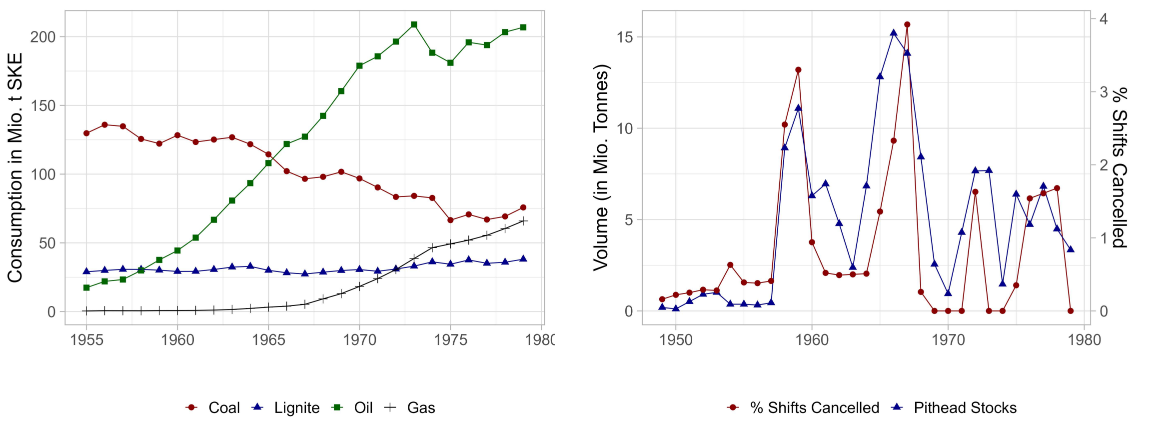
Historical Background. Coal mining has been a crucial component of the West German economy since the 19th century. Figure 4.1 illustrates the location of all post-WW II mines across the three coal districts *Aachen* (South West), *Ruhr* (center), and *Ibbenbüren/Lower Saxony* (North).

During WW II, coal mines were essential energy providers for German steel and arms factories, with output peaking in 1939 at the onset of the war. After the war, the industry quickly recovered from destroyed mines, fueled by the energy demand for the country's reconstruction.

In 1952, Germany joined the European Coal and Steel Community (ECSC), which abandoned tariffs and import restrictions between its member states Germany, France, Italy, Luxembourg, the Netherlands, and Belgium. However, even after joining the ECSC, intra-European coal trade remained relatively insignificant for Germany. During the 1950s and 1960s, coal imports (exports) accounted for 10 (15) million tonnes annually, representing only 8% (11%) of the output of German mines (see left panel of Figure 4.A.1 in the Appendix). Imports to Germany did not increase because German mines were more productive

than their main competitors, i.e., Belgian and French coal mines (see right panel of Figure 4.A.1).[3] Exports did not rise because mines produced at full capacity in the 1950s and due to the breakthrough of oil all over Europe from 1960 onwards.
Instead, coal imports from the US put pressure on the industry in the late 1950s as freight fees decreased by 80% (Bundestag, 1959*b*). In response, Germany introduced tariffs of 20 Deutsche Mark [DM]/tonne on non-ECSC coal beyond a tariff-free contingent of 5 million tonnes annually, effectively capping coal imports at this threshold (Bundestag, 1959*a*).[4]
By the late 1950s, coal mining accounted for 8% (3%) of the German industry (total) employment, 60% of energy production, and 6% of GDP (Federal Statistical Office of Germany, 1960).

Decline. In the 1950s, Germany began importing oil as a substitute for coal (Fritzsche and Wolf, 2023) due to lower oil prices and improved access (i.e., lower shipping costs, diplomatic relations with the Near East).[5] Even, the removal of a VAT exemption for heavy oil (Bundestag, 1960) did not slow down this trend. The left panel of Figure 4.2 shows the rise of oil in the German energy mix. The right panel illustrates the resulting overcapacities in coal mines. After 1957, excess coal had to be stored in pithead stocks (i.e., coal storages) and 3% of work shifts were cancelled.



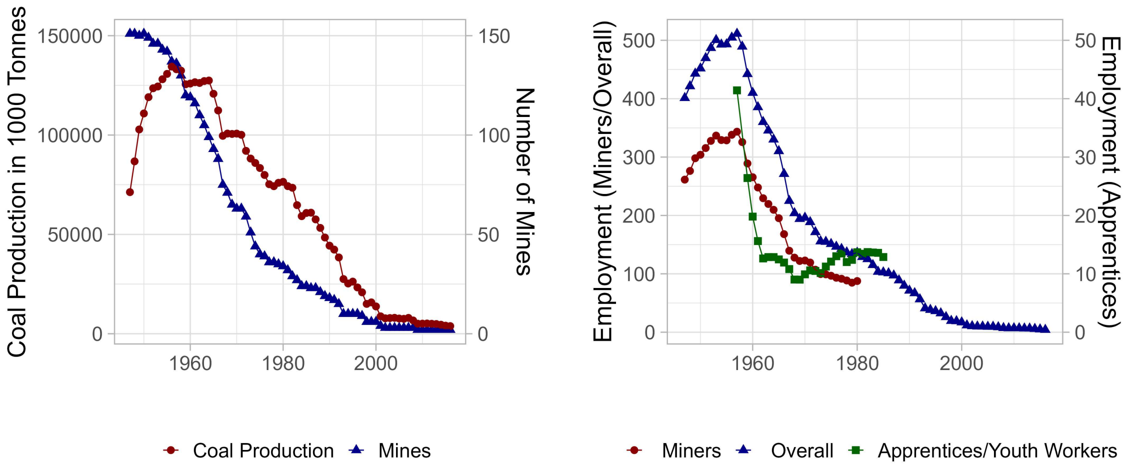
(a) German Energy Split (b) Measures of Overcapacities

Figure 4.2: Energy Split Transformation and Changes in Coal Demand

There was insufficient exit to eliminate overcapacities. Mines did not reduce their production, instead coal firms expected that more coal would be needed soon when the German economy experiences a new upswing (Bundestag, 1959*b*). Coal firms underestimated the impact of oil, predicting only a 2-3% decline in coal demand due to oil imports (Gatzka, 1996, Unternehmensverband Ruhrbergbau, 1961), while anticipating a 10-15% increase in coal pro-

[3] Also, coal prices were lower than in Belgium and France in the 1950s (see, e.g., High Authority (1956*c*)).
[4] Germany also introduced a one-year advertisement ban for oil in 1959 (Gatzka, 1996) and several other small but relatively ineffective measures (Stilz, 1969).
[5] To illustrate this substitution, in Table 4.B.1 of the Appendix, I show that the demand elasticities of coal with respect to the coal and oil price both increased in absolute terms after the uprise of oil. I instrument the coal price with average coal worker wages and the oil price with import wholesale prices from Saudi-Arabia.

duction by 1965 (Der Spiegel, 1958*b*). The High Authority of the ECSC, which oversaw conduct in the coal industry, even expected an increase in German coal production by 30% until 1975 in 1957 (Burckhardt, 1968).



(a) Production (Units)

(b) Employment in 1000s

Figure 4.3: Scope of the Industry Over Time

The first coal mines shut down in the early 1950s. Figure 4.3 shows the decline of the industry in terms of production and employment. From the early 1950s to 1970, the number of mines collapsed by 50%, output by 25%, total and miner employment by 60% and the number of apprentices by 75%. Figure 4.A.2 in the Appendix also shows the geographical distribution of the decline with earlier closures of small mines in the South of the *Ruhr* coal district.
The economic effects of mine closure were non-negligible given that 35% of all industry workers in counties with mine closures were employed in the coal industry in 1962 (Statistisches Landesamt Nordrhein-Westfalen, 1964). For example, on average a mine closure led to a persistent fall in the municipality (worker) population by 5-10% (Figure 4.A.3 in the Appendix).

Rationalisierungsverband des deutschen Steinkohlenbergbaus[6] (henceforth *RV*). This paper examines the consequences of the *RV*, a policy that became effective in 1963. The policy aimed at incentivizing the closures of inefficient mines and boosting the productivity of remaining mines (Bundestag, 1963). The main policy instrument was a closing premium of 25 DM per tonne of mine-level average production per annum between 1959 and 1961. The premium represented almost half of the market price. Half of the premium was paid by the government and half by the other firms in the coal industry. This premium structure was meant to capture the positive spillovers of closures on other firms. Only mines that would not run out of coal deposits shortly were allowed to take up the policy. However, all large mines met this criterion.
Given the urgent need for a capacity reduction around 1960, the policy was passed quickly, came unanticipated, and asked for soon closure decisions. In Figure 4.A.4 in the Appendix, I

[6]In English: Economization Union of the German Coal Mining Industry.

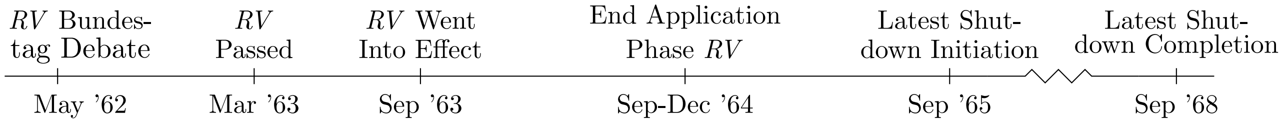


Figure 4.4: Timeline of the *RV*

show that those mines closed through the *RV* did not show anticipatory changes in output, employment, or productivity relative to non-exiting mines before the policy.

Figure 4.4 describes the policy rollout. After the law became effective in 1963, mine closures had to be announced within only one year and initiated within two years. By 1968, all mine closures had to be completed. The premium was only paid for coal fields that were permanently closed.

Mines that closed during the phase of the parliamentary debate but before the law became effective (May 1962 - September 1963) were also eligible, but only received 12.5 DM per tonne, the publicly financed premium. This upfront part of the policy was called *Vorausaktion*[7].

The policy closed mines with a pre-closure capacity of 31.5 million tonnes per annum, or around 25% of the overall industry output. Besides a few partial closures (i.e., only some coal fields of a mine were closed), the policy encompassed the closure of 23 large and 14 small mines. The *Vorausaktion* accounted for 8 out of the 31.5 million tonnes. Figure 4.5 plots the mine closure dates. Between 1962 and 1968, all closures were related to the *RV*. Few mines with only partial closures due to the *RV* closed after the 1960s. Overall, the policy scheme included payments of 590 million DM (in real terms of 2020: 1.5 billion Euro) - government and competitor payments combined.

The policy had two effects on workers illustrated in Figure 4.A.5 of the Appendix. First, the left panel shows that many workers were laid off, leading to a spike in the share of terminated contracts during the mid-1960s closure phase.[8] Using county-level employment data, I also show that there is a persistent decrease in the county-level number of coal workers in counties with closures due to this policy (right panel). There are no short-run spillovers to other industries. Second, workers who remained in the industry were partly shifted between mines. This resulted in a peak in the share of incoming workers who had previously worked at another mine.

Figure 4.6 compares how labour productivity[9] in the German and Belgian coal mining industries evolved around the *RV*. Belgium developed similarly to Germany before the policy and quarried from the identical cross-national coal field. Labour productivity increased more strongly after the *RV* in 1963 in Germany than in Belgium (on average 10% difference, in 1971: 12% difference).

[7] In English: Upfront payment.

[8] A high share of workers were sent to early retirement (German Federal Commissioner for Coal, 1968).

[9] Later on, I will use total factor productivity as the baseline measure of productivity. However, data availability only allows me to estimate total factor productivity for German mines, so that I compare Germany and Belgian in terms of labour productivity.

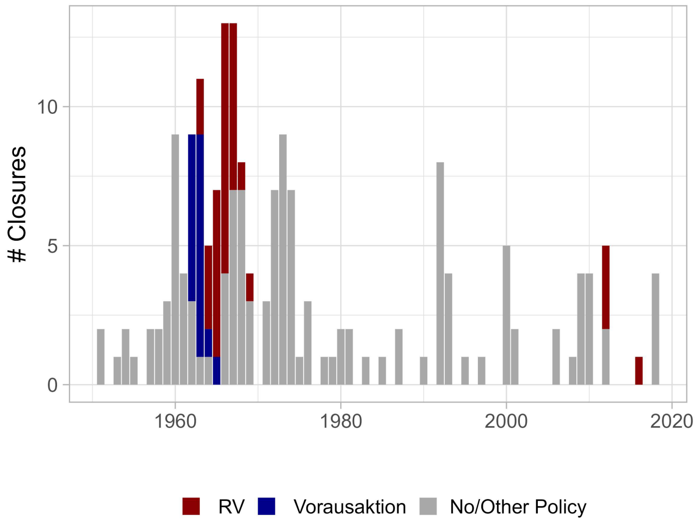


Figure 4.5: Closures by Supporting Policy

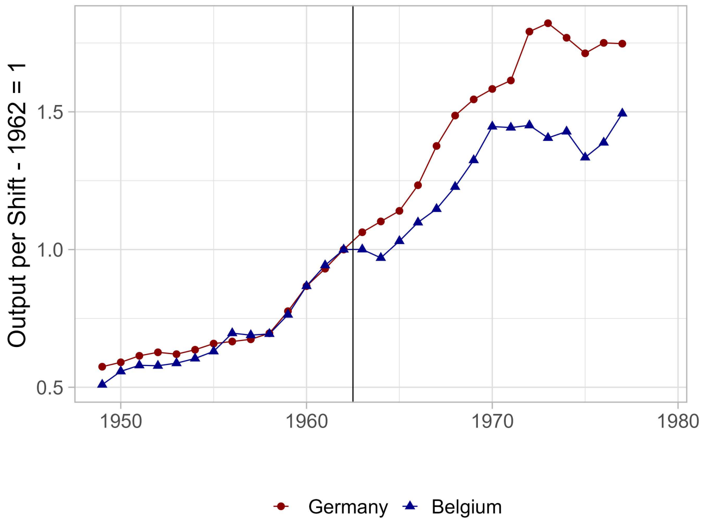


Figure 4.6: Comparison of Labour Productivity Development in Germany and Belgium

Price Setting and Conduct. All German coal firms sold their coal exclusively through three (after 1958: two) retail organizations, which they jointly owned.[10] These retail organizations were run as coal syndicates through which firms negotiated binding prices. Price deviations would have been heavily sanctioned due to the explicit contracts (Geitling, 1956). As the retail organizations should not make profits as determined in the contract, there was no double marginalization. The coal firms reported their expected output for the next quarter and year as well as broad expectations about the upcoming years (Geitling, 1960).

Even though each coal firm only belonged to one retail organization, actual competition between the retail organizations was doubted (Carret, 2023, Gatzka, 1996). They had their offices in the same building (High Authority, 1956*b*, Der Spiegel, 1963) and often announced (almost identical) price lists simultaneously (Gatzka, 1996, High Authority, 1965). Also, the

[10]Similar coal syndicates also existed in the Netherlands, Belgium, and France at the time (Der Spiegel, 1958*a*).

ECSC allowed them to collaborate to ensure the availability of the different coal types, to avoid harm to employees by smoothing demand across mines and to jointly save transportation costs (High Authority, 1956*a*,*b*).
Effectively, price setting in the retail organizations at the industry level left firms with only choosing their production output level as the main strategic variable.

4.3 Data

In this section, I present my data at hand and how I estimate total factor productivity, an essential variable to my analysis later on.

4.3.1 Mine-Level Production Information

I build a novel dataset on the universe of German mines from the 1950s to 1970s. For this, I digitized various data sources - primarily from the *Yearbooks of the German Coal Mining Industry*, the *Establishment Statistics* by the *Statistics of the Coal Industry e.V.*, and the *Annual Reports of the RV*. Other sources for mine-level geological information will be named below.

Sample Construction. For my analysis, I abstract from 'small' mines as well as mines from the Saar area. 'Small' mines are characterized by a very low output[11] and usually are run by industry companies or municipalities to exclusively serve their own consumption. Saar mines are excluded due to the region's unique post-war status as a French protectorate until 1957. Even after rejoining Germany, a significant portion of Saar's coal production was allocated to reparation payments to France, thus only partially contributing to the German coal market. The final, remaining sample encompasses about 90% of the German annual output, covering approximately 150 mines in 1952, with around 35 mines still active in 1980. I restrict my sample to data until 1971 as other policy interventions took place in the early 1970s, e.g. all coal firms were forced to merge into one entity.

Production Data. The dataset contains detailed mine-level annual input and output data. For output, the data includes both raw physical output (including non-coal content) and pure coal output in tonnes. For inputs, I observe four different inputs in physical units. First, for labour I observe aggregate employment, employment of below- and above-surface workers and administration workers as well as annual per-shift average wages. I can also track inflows of workers from other mines and from non-mining industries. Lastly, the number of cancelled shifts for various reasons such as accidents, production breakdowns, and a lack of demand is available.
Second, I observe capital, i.e., machinery power in kWh. Third, there is data on the electricity consumption in kWh. Fourth, the data includes pit wood consumption in cubic meters. Pit wood is used to construct and stabilize tunnels and as an energy source. I collect annual data on pit wood prices by cubic meters from Schroeder (1953) and the *Statistical Yearbooks of*

[11] The median output of a small (large) mine is 6.3 (1,135.3) tonnes per annum.

Germany.

Geological Mine-Level Data. Mine-level productivity is heavily determined by geological factors. For each mine and year, I know the maximum depth, the coal field size (i.e., the mining rights), and the number of seams (coal layers). Further, I observe the type of coal deposits a mine stores and quarries in its field as well as the average thickness and angle of the coal layers. Lastly, I gather data from the *Geological Office of Northrhine-Westphalia* on the thickness of the marl soil layer which is a sediment layer located between the surface and the coal layer. It mainly determines the depth of the coal deposits and set-up costs of the mine.

Technology Adoption. Data on the annual number of mining positions used for manual and mechanical mining as well as the share of production using these technologies is available.

Price Data. I know mines' prices through price lists of the retail organizations published in the *Statistical Yearbook of Germany* and the *Statistical Yearbooks of the German Coal Mining Industry.* I match prices to mines based on the 1963 allocation of firms to retail organizations. Prices changed multiple times a year, so that I calculate a time-weighted annual average price.
Mines sold various types of coal (e.g., fat coal, anthracite, gas coal) in different forms (e.g., cleaned/uncleaned, nugget/fine). However, due to limited differentiation in production, I treat mines as single-product establishments, following existing literature (Delabastita and Rubens, 2023, Rubens, 2023*a*,*c*). I employ coal prices for the common 'Nut III' form. The mine-year coal price then is a mine-level weighted average price of the mine-level production across coal types.

Firm-Level Financials. I collect firm-level stock values, dividends as well as assets and debts of stock market-listed firms from the annual *Salinger's Aktienführer.*

Descriptive Statistics. Table 4.B.2 in the Appendix summarizes all data.

4.3.2 Productivity Estimation

A main variable in my analysis is productivity. I will use two measures of productivity. First, labour productivity, i.e., output per worker shift, is a simple proxy for productivity, suits the labour-intensive production process in the industry well, and can be calculated directly from the data. Second, I estimate total factor productivity (TFP) to account for the consumption of all inputs beyond labour.

Production Function Estimation. I combine insights on structural production function and markup estimation by Ackerberg et al. (2015), De Loecker and Scott (2022), and De Loecker and Warzynski (2012). I start with the production function of a mine i in year t:

$$Q_{it} = \min\{\beta_{m,it} M_{it}, F(K_{it}, L_{it}, \beta)\Omega_{it}\} exp(\epsilon_{it}) \tag{4.1}$$

where Q_{it} is physical output in tonnes of coal which is produced with a production technology of capital K_{it}, labour L_{it}, and material input M_{it}. In my data, K_{it}, L_{it}, and M_{it} are given by the machine power (in kWh), worker shifts, and amount of pit wood (in cubic metres) - all in physical units. Ω_{it} is a Hicks-neutral productivity shock. The measurement error is given by ϵ_{it}.

I assume a Leontief production function where $F(K_{it}, L_{it}, \beta)$ is Cobb-Douglas with time-invariant output elasticities $\beta = \{\beta_l, \beta_k\}$ (see, e.g., Avignon and Guigue (2023), De Loecker and Scott (2022), Hahn (2024), or Rubens (2023*b*)), i.e., mines produce with a fixed ratio of pit wood and the combination of labour and capital. Pit wood, primarily used for stabilizing tunnels and as an energy source, is difficult to substitute with labour or capital. However, labour and capital can be substituted for each other.

Using physical input and output data avoids that estimation results are prone to input and output price biases (De Loecker and Scott, 2022). The Leontief production function avoids identification concerns for multiple flexible inputs (Gandhi et al., 2020) and also reduces concerns about unobserved conduct. Since coal is not a differentiated product, concerns about unaddressed quality biases are minimized.

I rely on standard timing assumptions of input choice (Ackerberg et al., 2015, Levinsohn and Petrin, 2003, Olley and Pakes, 1996) and assume that productivity follows a first-order Markov process:

$$\omega_{it} = g(\omega_{i,t-1}, [RV\ Exposure_j], Post_{i,t-1}, [RV\ Exposure_j] \times Post_{i,t-1}, Pr(Exit)_{it}) + \zeta_{it} \tag{4.2}$$

where $\omega_{it} = ln(\Omega_{it})$ is a function of its lagged value, the policy, the likelihood of market exit in the next year, $Pr(Exit)_{it}$, and an exogenous productivity shock (ζ_{it}). Policy exposure is given by (i) the share of pre-policy, firm-j-level production which has been closed due to the policy ($[RV\ Exposure_j]$), (ii) lagged values of a before-after policy dummy, $Post_{it}$, which turns one after the majority of exits from after 1965 onwards, and (iii) the variables' interaction.

First, I exploit that rearranging the logged production function gives an explicit control for productivity which identifies the measurement error ϵ_{it}:

$$\omega_{it} = ln(\Omega_{it}) = ln(\beta_{m,it}) + m_{it} - ln(F(K_{it}, L_{it}, \beta)) \tag{4.3}$$

We run a two-step estimation procedure. I first run an OLS regression of logged output on $ln(\beta_{m,it})$ and m_{it} where I approximate the former using a high-order polynomial of logged labour, capital, and materials (l_{it}, k_{it}, m_{it}), i.e., $\phi(.)$, as well as variables affecting input choices, X_{it}, such as mine depth and age, the existence of a cokery, number of coal layers, wages, and year fixed effects:

$$q_{it} = \phi(k_{it}, l_{it}, m_{it}, X_{it}) + \epsilon_{it} \tag{4.4}$$

This step provides an estimate for the predicted output, denoted as $\hat{\phi}_{it}$. The moment conditions, which identify the production function coefficients $\beta = \{\beta_l, \beta_k\}$, are given by:

$$E[\zeta_{it}(\beta) \begin{bmatrix} l_{i,t-1} \\ k_{it} \end{bmatrix}] = 0 \tag{4.5}$$

Timing assumptions denote that capital is chosen before labour. Labour has been quite flexibly adjustable. For example, in response to the policy 87,000 jobs were cancelled or shifted between mines immediately. An estimate for ω_{it} is then calculated by:

$$\hat{\omega}_{it} = \hat{\phi}_{it} - ln(F(K_{it}, L_{it}, \hat{\beta})) \tag{4.6}$$

Our baseline TFP estimates yield a correlation of 0.795 with labour productivity.
For markup estimation, I rely on De Loecker and Warzynski (2012) and De Loecker and Scott (2022). Cost minimization with respect to input choices gives markups $\hat{\mu}_{it}$ and marginal costs $\hat{c}_{it}$:

$$\hat{\mu}_{it} = \frac{P_{it}}{\hat{c}_{it}} = \frac{1}{\frac{\eta^L_{it}}{\hat{\beta}_l} + \eta^M_{it}} \qquad \hat{c}_{it} = \frac{P_{it}}{\hat{\mu}_{it}} \tag{4.7}$$

where P_{it} is the per-tonne price for coal and η^L and η^M are the revenue shares of labour and materials expenditures, corrected for the measurement error. Table 4.B.3 in the Appendix provides the estimation results, i.e., output elasticities, scale parameters and markups, for the baseline approach and a robustness check with a Cobb-Douglas production function with electricity as substitutable material input. The production is labour-intensive and the scale parameter is not significantly different from one. To ensure that my results are not sensitive to strong assumptions on the production technology, I will provide several robustness checks in Section 4.6.

4.4 Empirical Strategy

To examine *RV*'s effect on industry-wide productivity, I study mine-level outcomes along various margins. Figure 4.6 showed that labour productivity in German relative to Belgian mines grew due to the policy but the underlying channels remain unclear. My empirical strategy is twofold. First, recall that the policy intended to push out unproductive mines. I study this extensive margin effect of the policy by examining whether productivity actually is a determinant of exit through the policy. Second, I study potential (un)intended side effects of the policy. I look at the mines' changes in production at the intensive margin, i.e., whether market shares are reallocated to more productive mines and whether firms use the policy to improve productivity. I explain all steps below.

Extensive Margin. From an ex-ante perspective, it is unclear whether exit is negatively correlated with productivity, which was the main policy goal. Admittedly, unproductive mines with higher marginal costs are more likely to be closed ceteris paribus. However, firms will also consider output spillovers from closing a mine on their remaining mines, which are

heterogeneous across mines. For example, multi-mine firms internalize some spillovers in their remaining mines if they shut down a mine. Single-mine firms cannot do so as they close their only mine. Also, the costs of closing a mine could be very heterogeneous and potentially higher for machinery-heavy, productive mines.
To examine the relation between mine-level closures and productivity, I run a cross-sectional regression of a closure dummy on mine-level productivity and controls:

$$1[Closure\ via\ RV]_i = \alpha\ (+\ \lambda_j) + \gamma_d + \theta Prod_i + X_i'\zeta + \epsilon_i \tag{4.8}$$

where $1[Closure\ via\ RV]_i$ is a dummy turning one if a mine i closed through the *RV*. The constant is given by α and $Prod_i$ is a measure of mine i's productivity right before the policy (1959-1961), either labour productivity or TFP. I include coal-district fixed effects, γ_d, for the districts *Aachen*, *Ruhr*, and *Ibbenbüren/Lower Saxony* to account for region-specific drivers of exit. In some regressions, I add firm fixed effects, λ_j, to distinguish between across- and within-firm variation.
A simple OLS regression likely yields biased results. First, expectations about a mine closure can affect productivity shortly before the closure (*reverse causality*). Second, mine-level productivity is shaped by unobserved factors that also impact the closing decision (*omitted variable bias*), e.g., whether a mine has refinement plants (e.g., power plant) attached, the economic potential of a region, as well as local differences in industrial policy. Ex-ante, the direction of the OLS bias is unclear. A better future economic potential would likely imply higher productivity and could ease structural change, so that exit is not postponed ($\hat{\beta}_{OLS} > \beta$). Good policymaking could increase mine-level productivity and delay exit ($\hat{\beta}_{OLS} < \beta$).
I address the endogeneity problem by means of an instrumental variable approach. A mine-level instrument needs to be a relevant shifter of productivity, should not correlate with unobserved drivers of the exit decision, and should not have a direct effect on closures. I use geological conditions which affect mine productivity as an instrument. Geology, i.e., the nature of below-surface sediment layers, however, should not directly affect business leaders' decision of whether to close a mine or not (the outcome variable). I rely on three geological measures of coal degradability - illustrated in Figure 4.7 - which affect mine-level productivity. First, mines' *coal angle* (see box $\boxed{1.}$ in Figure 4.7), i.e., the coal layer's steepness. The higher the coal angle, the more difficult machinery usage and construction work of tunnels. I use the share of coal deposits with a coal angle of up to 25 degrees.
Second, the depth of the coal layer, the so-called *marl thickness* $\boxed{2.}$. Marl is the sediment layer between the surface and the coal layer. The thicker this layer, the higher the set-up costs of a mine. Mines, therefore, were only profitable to build in thick-marl regions when they were more productive to break even. Also, the necessary technology to break through the marl layer was only available by mid-19th century, so that mines in thick-marl regions typically are younger with wider tunnels. That makes the adoption of large-scale machinery more feasible. Third, the *seam thickness* $\boxed{3.}$. Thicker coal layers give a higher return on machinery usage.
In Panel (a)-(c) of Figure 4.A.9 in the Appendix, I show that all three measures strongly

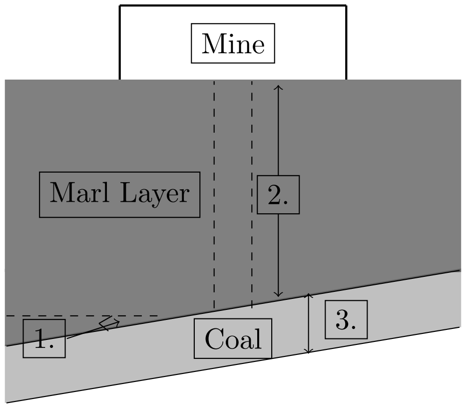


Figure 4.7: Sketch of a Mine and Coal Degradability Factors

correlate with labour productivity (coal output per worker shift) and TFP. To combine all three measures into one IV, I conduct a principal component analysis which joins the isolated, independent variation from all three variables. The constructed variable is a strong predictor of productivity (panel (d)).

By assumption, the IV needs to satisfy random assignment holding mine characteristics fixed (i.e., conditional independence) and should only affect exit through the productivity channel (i.e., exclusion restriction). Conditional independence is ensured by controlling for the main sources of mine heterogeneity in exit decisions and not including any kind of mine-specific factors beyond geology as part of the instrument. With regard to the exclusion restriction, the IV should not affect exit decisions beyond shifting mine-level productivity. If the geology affected, for example, not-included local industry structure, this would invalidate the IV. Figure 4.A.10 in the Appendix provides conditional correlations of the IV with mine-specific measures of ownership, vertical relations, transportation networks as well as local economic strength and industry composition. The IV is not significantly correlated with these measures. Lastly, there is no reverse effect of productivity on geology, i.e. firms of heterogeneous productivity do not select into different geology. Mines have mainly been established in the 19th century when vertically integrated industry firms primarily chose to quarry for coal in the nearest possible coal field. Also, geology as a main driver of machinery productivity has been a less important productivity determinant in the times of labour-intensive work back then (Gebhardt, 1957).

Intensive Margin. To analyze how firms adapt their production after the RV, I examine the RV's effects on mine-level outcomes in a difference-in-differences setup. I compare surviving mines of firms, that heavily took up the policy, to less-exposed firms and non-uptakers, before and after the policy. I estimate a dynamic, continuous exposure difference-in-differences regression:

$$Y_{it} = \alpha_i + \beta_{dt} + \sum_{\tau, \tau \neq 1962} \delta_\tau [RV\ Exposure]_j \times 1[Year = \tau]_t + u_{it} \tag{4.9}$$

where Y_{it} is an outcome of mine i owned by firm j in year t. Mine and coal district-year fixed effects are given by α_i and β_{dt}. $[RV\ Exposure]_j$ is a treatment variable that is the share of pre-policy capacity which owner j of mine i closed through the *RV*. If a firm shuts down all its mines, the variable will be 1. If it did not shut down a single mine, it is 0. While firms actively decide on their exposure level, the timing and size of this shock is exogenous from the perspective of a surviving mine. Accounting for the rich variation in the uptake and exposure to the policy in a continuous treatment variable is more precise than a binary treatment of policy uptake or not. τ mostly ranges from 1956 to 1971 but the window can narrow for data availability reasons for some outcomes. To only capture spillovers to the unaffected mines, I do not include the closed mines in the regressions. As they were only operating before the policy-induced closures, they would not contribute to the identification of changes from before to after the policy anyway.

I identify a causal effect of the policy uptake under two assumptions. First, the parallel trends assumption implies that mines of firms with different exposure levels would have evolved similarly absent the policy. I will provide suggestive evidence for this by looking at the pre-trends. This also is an implicit test for potential reverse causality. In particular, if changes in mine outcomes led to the uptake of the policy, this would cause an upward or downward pre-trend.

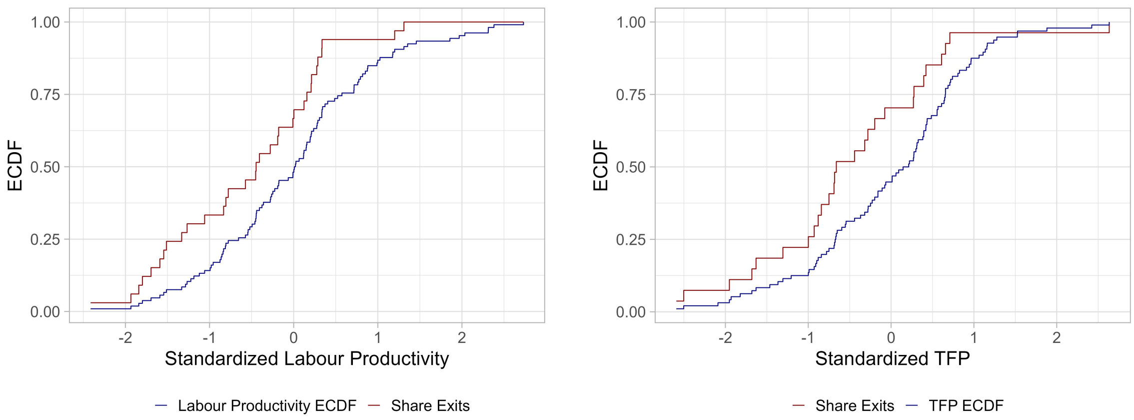
Second, the stable unit treatment variable assumption (SUTVA) has to hold. It requires that the outcome and treatment status of one mine is unaffected by the treatment status of other mines. In my setting, however, it could be the case that there are spatial spillovers between treated and untreated mines, so that treatment effects should be interpreted as relative effects between treated and untreated mines. For example, the closure of a mine affects non-treated mines through a potential inflow of workers. We later on provide insights that spillovers across firms in space are limited mitigating the SUTVA concerns.

In the last step, I further explicitly estimate the effect of the policy on misallocation as reallocation could explain productivity changes in the industry, too. Misallocation is usually identified by showing that there is a negative or only weakly positive correlation between productivity and market share (Baily et al., 1992, Bartelsman et al., 2013, Griliches and Regev, 1995, Melitz and Polanec, 2015, Olley and Pakes, 1996, Pavcnik, 2002). In the triple-difference analysis, I can analyse how the policy affects the relation between market share and productivity in treated firms:

$$s_{it} = \alpha_i + \beta_{dt} + \sum_{\tau,\tau\neq 1962} \delta_\tau Prod_{it} \times [RV\ Exposure]_j \times 1[Year = \tau]_t + \Gamma'_{it}\zeta + \epsilon_{it} \tag{4.10}$$

where s_{it} is mine i's market share in period t. Γ_{it} is a matrix including the further variables and interactions of the triple-difference model. An increasing relation between productivity and market share would be expressed by positive values for δ_τ if $\tau > 1962$.

Selection into Treatment. Mines differ in their characteristics which impact their owners' decision to exit or not. Hence, treated and non-treated mines of different $[RV\ Exposure]_j$ might structurally differ. I provide a comparison of both groups in Table 4.B.4 in the Ap-



(a) Labour Productivity

(b) TFP

Figure 4.8: Empirical Distribution Function of Productivity and Policy Uptake (Exit)

pendix. There, I regress the treatment status on a number of mine-level variables. I show that there are almost no differences. The only robust disparity is that treated mines belong to bigger mining firms. This is by construction as only mines of multi-mine firms can be treated because at least one mine of the same firm has to exit. Hence, I cautiously interpret these results as in favor of the treated mines being close to representative of all surviving mines. I later on provide robustness checks that my main results also hold for a subsample of only multi-mine firms.

4.5 Results

Recall that we are interested in the productivity effects of the policy. In this section, I provide results on three channels via which the policy affects industry-wide productivity: First, indeed, the policy closes especially low-productivity mines. Second, firms reallocate production of closed mines to more productive, remaining mines. Third, firms increase within-mine productivity of remaining mines after closing one of their other mines. I consecutively discuss these channels in the subsections below. I also mention underlying mechanisms and stock market responses.

4.5.1 The Effect on Exit

First, I test the policy's attempt to force unproductive mines out of the market. Figure 4.8 descriptively plots the empirical distribution functions of productivity (blue) against the cumulative policy uptake, i.e., exit, across mines (red). Productivity first-order stochastically dominates cumulative exit. Hence, there is a negative correlation between productivity and exit in the sample.

Table 4.1 investigates the relationship between exit and mine-level productivity. Panel A provides the OLS and Panel B the IV results.

Panel A reveals a significant negative relationship between exit and both productivity measures. A one standard deviation increase in labour productivity or TFP implies a reduction in the likelihood of exit by 14 and 23 p.p. respectively. This is robust across (columns (1) and (4)) and within firms (i.e., including firm fixed effects, columns (2) and (5)). Lastly, in columns (3) and (6), I control for additional variables likely affecting exit. I highlight two of the variables in the regression table: (i) a dummy indicating that a mine had an above-median share of cancelled shifts due to insufficient demand and (ii) a dummy indicating that a firm has closed a mine which produces the same coal type as mine i through the policy. Insufficient demand is a driver of exit and the policy uptake becomes less likely if a firm already closed another mine producing the same type of coal. The relation between productivity and exit is unaffected by the inclusion of the controls.[12]

In Panel B, I then examine the causal relationship using the preferred IV model. I instrument productivity with the 'coal degradability' IV including the information about a mine's geology. The first stage F-Statistic is above the threshold of 10 in all specifications, i.e., the IV is relevant. The IV results show a negative causal effect of productivity on exit about twice as large as the OLS coefficients. Additionally, I provide results on marginal costs (based on the production function approach) instead of labour productivity or TFP in Table 4.B.5 in the Appendix. Marginal costs also take into consideration input prices and not just the efficiency of inputs. The table also provides the IV results when using the three geology dimensions as IV individually.

To judge the effect of this positive selection of mines, I compare the average, weighted labour productivity of the remaining mines with the overall sample of mines right before the policy uptake. I find that the remaining mines have a 17.4% (7.9%) higher labour productivity (TFP). Exit therefore increases industry-level productivity by 3.1% (1.5%).

I assess how close the observed selection is to an optimal selection benchmark, i.e., exit ordered by productivity rank. I calculate the productivity gain from the exit of the least efficient mines which sum up to the same exit output volume as actually observed. This exit, for example, would have increased industry-wide productivity by 4.8% (3.3%) with respect to labour productivity (TFP). Hence, the observed selection worked arguably well given that productivity is not the only factor driving the exit decision.

Finally, the finding of relatively efficient exit orders adds to the empirical literature on exit orders and industry shakeout (Gibson and Harris, 1996, Hünermund et al., 2015, Klepper and Simons, 2000, 2005, Klepper and Thompson, 2006, Lieberman, 1990, Takahashi, 2015). Theoretical work has shown that exit in multi-entity firm environments might be inefficient as smaller firms can commit to staying in the market given a decreasing demand for a longer time (Fudenberg and Tirole, 1986, Ghemawat and Nalebuff, 1985, 1990, Whinston, 1988).

4.5.2 Output Reallocation

Output Spillovers Within-Firm. Second, the policy could have affected industry productivity through other channels beyond exit. In particular, firms could have reallocated their production across their mines after a mine closure. I first look at how the policy uptake affected the output of a firm's remaining mines by estimating equation (4.9) with logged

[12]Also including further controls such as the mine age does not affect the results.

Table 4.1: Selection of Plants into *RV* by Productivity

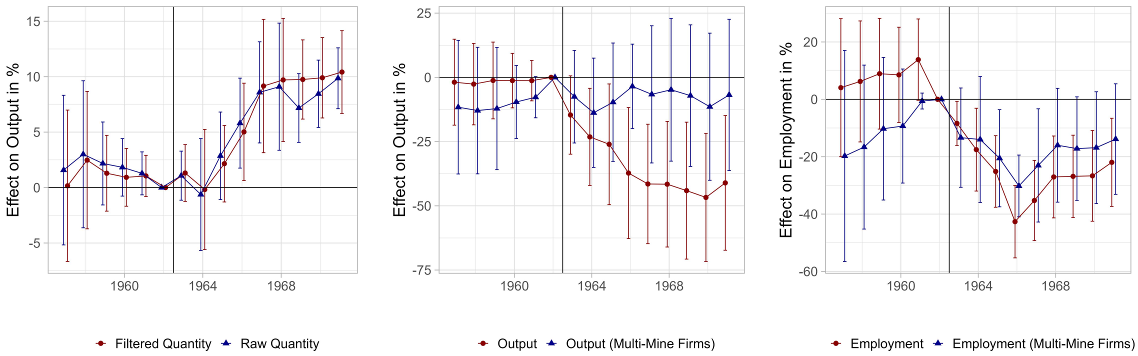
	1[Closure via RV]$_i$					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Standardized LP$_i$	-0.183^{***}	-0.210^{***}	-0.198^{***}			
	(0.033)	(0.036)	(0.042)			
Standardized TFP$_i$				-0.155^{***}	-0.261^{***}	-0.206^{***}
				(0.055)	(0.068)	(0.067)
1[High Cancelled Shifts]$_i$			0.458^{***}			0.384^{**}
			(0.125)			(0.146)
1[Closure of Same Coal Type Mine]$_i$			-0.513^{***}			-0.379^{***}
			(0.085)			(0.101)
Panel B: Instrumental Variable						
Standardized LP$_i$	-0.528^{***}	-0.483^{***}	-0.358^{***}			
	(0.162)	(0.139)	(0.129)			
Standardized TFP$_i$				-0.287^{***}	-0.409^{***}	-0.358^{***}
				(0.065)	(0.109)	(0.129)
1[High Cancelled Shifts]$_i$			0.461^{***}			0.331^{**}
			(0.116)			(0.164)
1[Closure of Same Coal Type Mine]$_i$			-0.470^{***}			-0.243
			(0.133)			(0.155)
F-Statistic First Stage	*10.63*	*12.46*	*10.96*	*50.78*	*33.56*	*24.89*
Mining District FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm/Owner FE	No	Yes	Yes	No	Yes	Yes
Observations	106	106	97	96	96	96

Note: Significance levels of 10%, 5% and 1% are denoted by *, ** and ***. Labour productivity is averaged over 1959-1961. TFP is averaged over 1960-1961 (no data for 1959). Standard errors are clustered at the firm/owner level.

output as the outcome. The left panel of Figure 4.9 shows that the output of mines with mean treatment exposure expands by around 10% (coefficients are multiplied with the mean [$RV\ Exposure_j$] of mines with positive exposure) relative to mines of non-treated firms. Hence, firms close mines through the policy but increase production in the remaining mines. As increasing production needs improved infrastructure beforehand and given that not all exits were immediately in 1963, these spillovers occurred a few years after the policy. To prove that the output growth is not driven by an increase in coal capacity or mergers and acquisitions, I show that the coal field size of mines and the probability of a coal field merger do not change with the policy (see left panel of Figure 4.A.7 in the Appendix).[13] Instead, I show that more work is done within the same mine borders. For example, more seams, i.e., coal layers, are worked on within the same coal field (right panel).

To understand the relative size of the output spillovers in remaining mines in comparison to the closed capacities, I run an analysis on output at the firm level before and after the policy (middle panel of Figure 4.9), i.e., a difference-in-differences regression in the style of equation (4.9), where I interact [$RV\ Exposure_j$] with year fixed effects conditional on firm

[13]Figure 4.A.8 in the Appendix shows how coal fields are distributed between coal firms as of 1962.



(a) Mine-Level Output (b) Firm-Level Output (c) Firm-Level Worker

Figure 4.9: Existence and Scale of Spillovers Within Uptaker Firms

Note: Left panel is based on equation (4.9). Middle and right panels based on equation (4.9), too, however at the firm level. For left panel: Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Regressions in middle and right panels weighted by 1961 firm-level output to account for firm-size differences. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

and coal district-year fixed effects. Firm-level output is normalized to 100 in 1962. For a one percentage point increase in the exposure, i.e., a coal firm officially closed one p.p. of the pre-policy output and earned the respective subsidies, a coal firm effectively only reduces output by 0.34 p.p. on average over the period 1963 to 1971 with a maximum effect size of 0.45 p.p. right after the policy. Thus, more than half of the closed volume is recovered in the remaining mines of treated firms. Firms earn an effective closure subsidy of more than twice the official 25 DM/tonne. That the effect is not larger than 0.45 p.p. even right after the policy indicates that firms shift mines and output to other mines very soon after the closure. I then look at a subsample of only multi-mine firms (red line). Single-mine firms by construction cannot shift production to other mines of the same firm after a closure. I find that multi-mine firms do not decrease their production volume at all.

Doing the same for employment instead of output, I show that per 1 p.p. of closed capacity, the net employment loss is only 0.33 p.p., supporting the spillovers described above. Again, multi-mine firms show a smaller but significant effect. The unchanged output volume next to lower employment already hints at potential gains in productivity in treated firms.

Heterogeneity in Mine-Level Output Reallocation. For policymakers, it is essential to understand what triggers output reallocation. I perform heterogeneity analyses along multiple mine dimensions. I subsequently show the policy induced reallocation to large, efficient mines.

First, I study whether mines with high productivity have the strongest output increases due to the policy. The first panel of Figure 4.10 shows that, indeed, mines that had an above-median labour productivity or TFP in the pre-policy year 1962 face the strongest output increase. Above-median mines of treated firms are 31% (16%) more productive in terms of labour productivity (TFP) than closing mines. Given that about half of the closed capacity (i.e., about 10% of the industry production) is reallocated towards these productive mines (see Figure 4.9), this reallocation increases industry-wide labour productivity (TFP) by 3.2% (1.7%).

Second, I show that output increases especially in mines that had some form of coal refinement plant attached, i.e., either a cokery or an electricity plant. Cokeries and electricity plants were stable consumers of coal, in contrast to declining coal demand for household and industry consumption. This indicates that the policy led toward a shift of the business model along the value chain and to reallocation to more secure and less volatile demand segments. Beyond that, I investigate further dimensions of heterogeneity to explain the underlying reasons for reallocation. Reallocation takes place towards larger mines (in terms of pre-policy output) which on average are also more productive (see Figure 4.A.11 in the Appendix). I further show that spillovers are especially strong for mines that experience exit of mines from the same firm nearby (2.5 or 5km radius). Hence, the geographical distance matters[14]. Also, output expansions are especially strong in larger firms with a larger number of mines.

Lastly, I show that output increases are largest for mines with a higher number of worker flats per capita, which have been operating for the shortest time, and with a low share of apprentices right before the policy introduction. Housing availability is necessary to be able to increase the number of workers. Younger mines on average have a higher mechanization rate. A lower share of apprentices implies a higher input quality, so that reallocation takes place to mines with better labour input. But also, the costs of laying off workers increase with age, so that reallocation towards mines with a lower share of apprentices can also be explained by the higher opportunity costs of not shifting capacities to such mines. All these heterogeneity analyses prove that reallocation takes place to productive, large mines with substantial potential to increase production further.

As a second piece of evidence that reallocation towards more productive mines took place, I conduct an analysis motivated by productivity decompositions in the fashion of Olley and Pakes (1996) and others. These papers argue a reduction in misallocation is achieved when the covariance between market share and productivity increases, i.e., more productive firms produce more in relative terms. I convert this intuition to an empirical test in a difference-in-differences framework following equation (4.10). Figure 4.11 shows that after the policy, indeed, the relationship between productivity and market share became stronger. A one standard deviation increase in productivity increases the market share of a mine with an average exposure by 0.25 p.p. (or 15%) after the policy.

Adjustments in Input Decisions. The increase in output raises the question of how input choices changed in policy-uptaking firms after the policy. For this, I study labour and capital. Figure 4.12 presents the policy uptake's effect on the remaining mines of a treated

[14] As an underlying channel, I document that especially worker flows from closing mines to surviving mines take place in response to exit. Figure 4.A.12 in the Appendix shows that mines that had exit from a different mine of the same firm nearby experienced an inflow of educated workers. The share of experienced workers from closed mines coming over to the surviving mine nearby (2.5km radius) increases by 40 p.p. for three years after the policy. Mines with closures farther away do not experience an increase in educated workers delivering another reason for fewer spillovers to such mines.

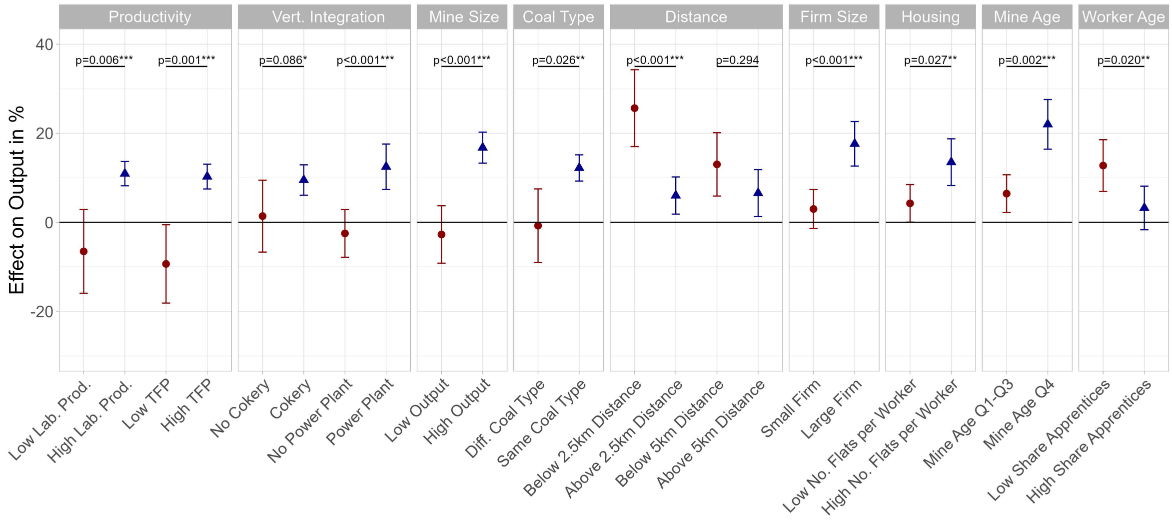


Figure 4.10: Reallocation along Dimensions of Heterogeneity

Note: Estimates come from a triple-differences estimation based on equation (4.9) where I pool all post-treatment years after the majority of exits occurred, i.e., after 1965. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

firm. In line with the documented output increase in treated mines, I find that capital, as well as employment, rise strongly after the policy. The power of machinery used in the mine increased by up to 20% until 1971 for a mine of average policy exposure (see left panel of Figure 4.12). Also, employment of miners as well as non-miner employment rose by up to 8% a few years post-policy (see right panel of Figure 4.12). I also show that the capital intensity, i.e., the machinery power per worker in the mine, increases significantly by up to 10%, so that the policy led to a more machinery-intensive production process.

4.5.3 The *RV*'s Effect on Dimensions of Productivity

I showed that the policy leads to exit of unproductive mines and to reallocation towards large, productive mines. These changes in the industry competition raised average productivity. Now, I will further show that the policy fostered within-mine productivity growth.

Productivity, Markups, and Marginal Costs. The policy targeted increasing industry-wide productivity through the exit of unproductive mines. Productivity gains can also stem from within-mine productivity growth which I examine subsequently. The left panel of Figure 4.13 shows a strong and persistent increase in labour productivity (TFP) after the policy in mines of policy-uptaking firms by up to 0.2 (0.3) standard deviations or 6.0% (7.3%) respectively for the mean exposure. For TFP, I present results based on two production function estimations as described in Section 4.4 - a Cob-Douglas production function with electricity (data until 1969 only) and a Leontief production function with pit wood. Hence, the policy led to strong productivity increases within mines that have sibling mines closed through the policy. Since not all mine closures occurred immediately in 1963, it is intuitive that the effects

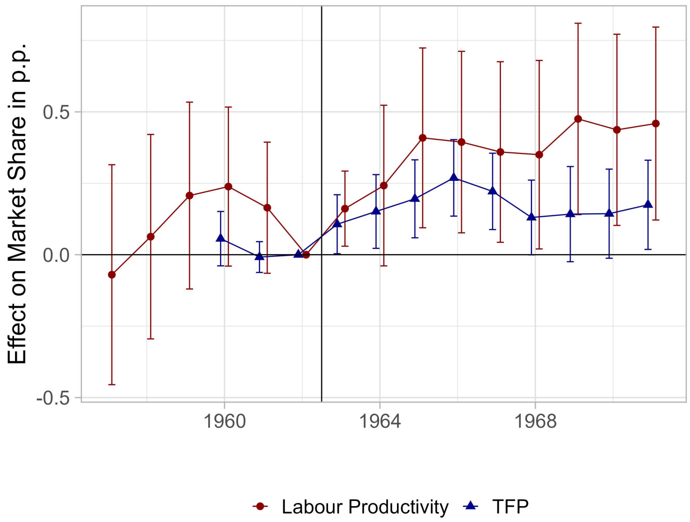


Figure 4.11: Effect on Reallocation

Note: Based on equation (4.10). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. This plot documents how the policy affects the relation between productivity and market share. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

manifest gradually over time.

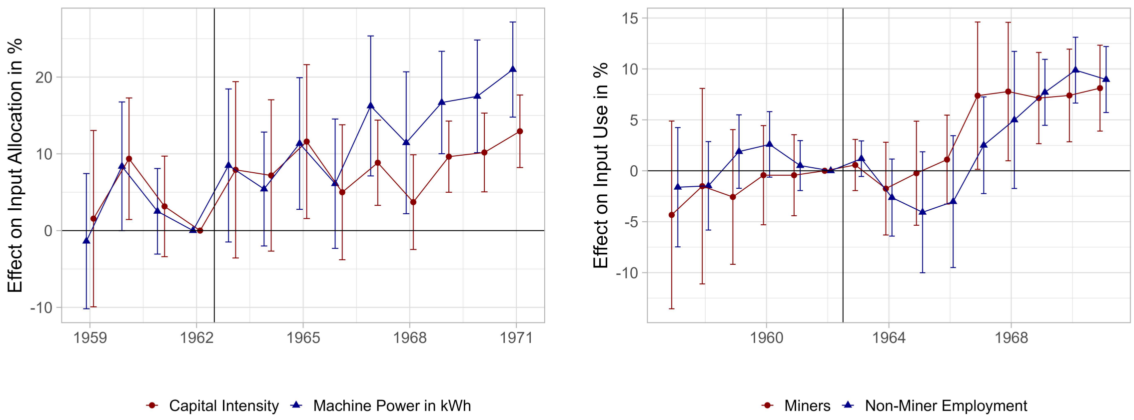
Note that the treated mines made up about half of the production in 1971, so that the 6.0% (7.3%) increase in labour productivity (TFP) translated to an about 3% (4%) increase in labour productivity (TFP) in the industry.

Productivity gains can come from investments or economies of scale. To show that these productivity gains are not just caused by potential economies of scale, i.e., the increase in output as shown in Figure 4.9, Figure 4.A.13 in the Appendix controls for mine-level output as a 'bad control'. The treatment effect on productivity is only partially explained by output changes, motivating that investments could play a crucial role for the arising productivity gains. I look at this mechanism later on in Section 4.5.4.

I further examine how the policy affected marginal costs and markups. The right panel of Figure 4.13 shows that marginal costs drop in response to the policy - at least temporarily. Marginal costs on average decrease by 1.5% which translated to cost savings of approximately 400 million DM over the nine post-effect window years. Hence, costs savings exceed the overall premium payments which the state paid through this policy. As expected due to the price setting in the retail organizations, prices do not change with the treatment.

Interestingly, TFP increases in the long-run but marginal cost savings are only temporary. As I will show later on, this is driven by increasing mine-level wages in treated mines in the long-run, which I use for the markup calculation as labour is the variable input - see point 'Markdown' in Section 4.6. There, I also show that when using electricity as material input, marginal cost gains seem to be more persistent over time.

Survival. Given that mines increase productivity, I examine whether these investments lead



(a) Capital

(b) Labour

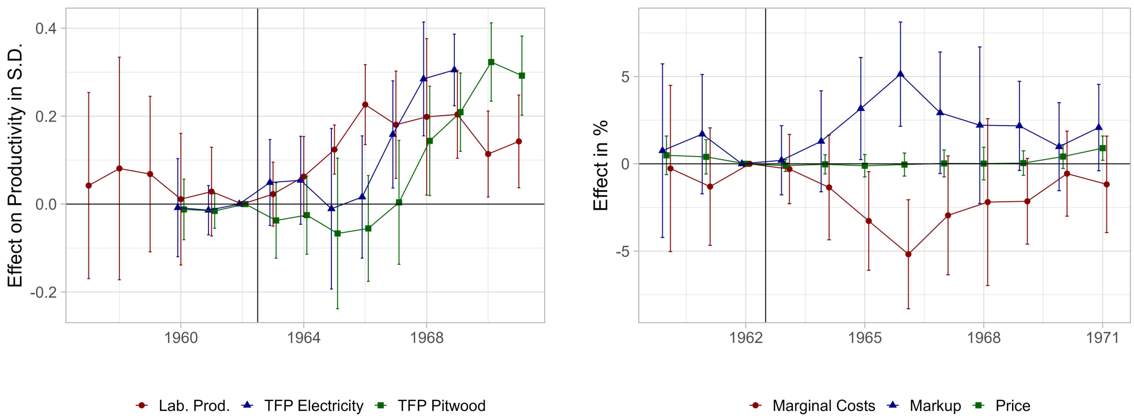
Figure 4.12: Effect on Input Consumption

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. These plots document how the policy uptake affects the input allocation of remaining mines of treated firms. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

to longer survival. Figure 4.14 plots the effect of the policy on the likelihood of not having exited yet. Exposed mines are on average 10 p.p. or 21% more likely to survive for three decades post-policy. For mines that are treated with more than the median $[RV\ Exposure_j]$, the effect is even stronger (20 p.p. or 42%) and significant for longer. The increase of 10 p.p. translates to a lifespan extension of 5.7 years for policy-uptaking mines.

Productivity Dispersion. Given that coal firms used the subsidy to reallocate production to their largest, most efficient mines, it is an open question what happens to their weaker mines. To unveil heterogeneous effects on productivity for mines with different ex-ante productivity level, I estimate (unconditional) quantile treatment effects of the policy along the productivity distribution in a similar fashion to Chen et al. (2022). I estimate the counterfactual distribution of productivity based on the distribution regression method by Chernozhukov et al. (2013). It obtains the counterfactual distribution (i.e., productivity of treated mines absent treatment) by estimating the pooled version of the difference-in-differences equation (4.9) with an adapted outcome variable. It estimates the effect of the policy exposure on a dummy that will be one if a mine-year observation has a productivity below a certain cutoff value. Adding the estimated treatment effect to the empirical distribution function of observed productivity then gives the value of the counterfactual distribution at this particular cutoff value. Repeating this for many cutoffs gives the full counterfactual distribution function. The first two panels of Figure 4.15 show that the within-firm spillovers are associated with productivity increases at the upper tail of the productivity distribution and a negative change at the left tail.[15] Hence, already productive mines become even more productive and unproductive mines deteriorate relative to more productive mines. Impor-

[15] This fits the distributional effects in Behrens et al. (2020) who show for many sectors that unproductive firms produce too much from a welfare point of view.



(a) Productivity

(b) Markups and Marginal Costs

Figure 4.13: Effect on Mine-Level Productivity, Markups and Marginal Costs

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

tantly, this productivity dispersion is based on the sample of surviving mines. As shown in the right panel, this is driven by the fact that the policy leads to output increases in very large mines. As Syverson (2004, 2011) argues for revenue-based TFP, productivity dispersion correlates with increasing market concentration and is a measure of inefficiency. A similar argument holds for physical TFP as I use in my analysis.

With regard to reallocation, panel (c) shows that the policy led to a growth in mine size of formerly already very large mines. This is further support for reallocation towards large, efficient mines. Hence, the policy set up very large, productive mines endogenously. Absent the policy, the largest mine of treated firms would have been smaller than one fourth of the treated mines after the policy. Figure 4.A.14 in the Appendix runs the distributional analysis for mine-year market shares and shows that the policy-induced reallocation was a main driver for the existence of mines with more than 2% market share. The policy more than doubles the share of mines with at least 2% market share.

4.5.4 Mechanisms

In this subsection, I study the underlying drivers of reallocation and productivity gains. I find that the subsidies earned alleviate financial constraints. That allows firms to invest in their infrastructure and technology adoption.

Financial Constraints. While I showed that output spillovers and productivity growth took place in response to the policy, it remains unclear what exactly triggered this transformation. Is it that mines that receive more closing payments or are more financially constrained before the policy respond more strongly (*financial constraints*) or is it that the found effects stem from other mechanisms such as general spatial, across-firm spillovers from mine closures (*local spillovers*)? I examine these channels subsequently.

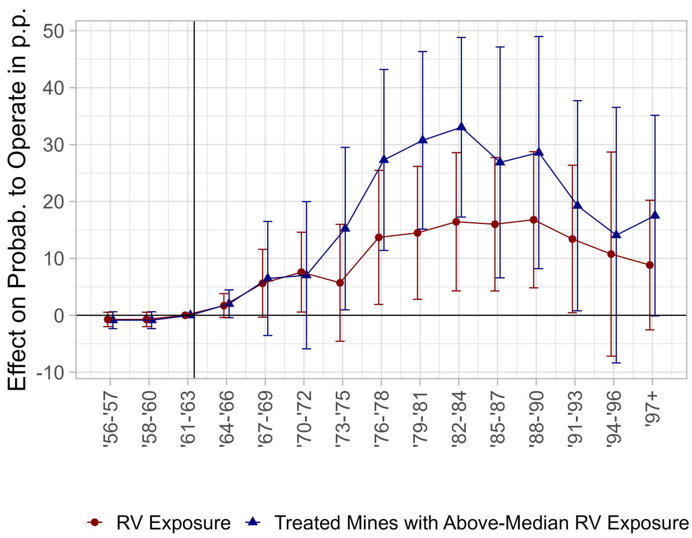
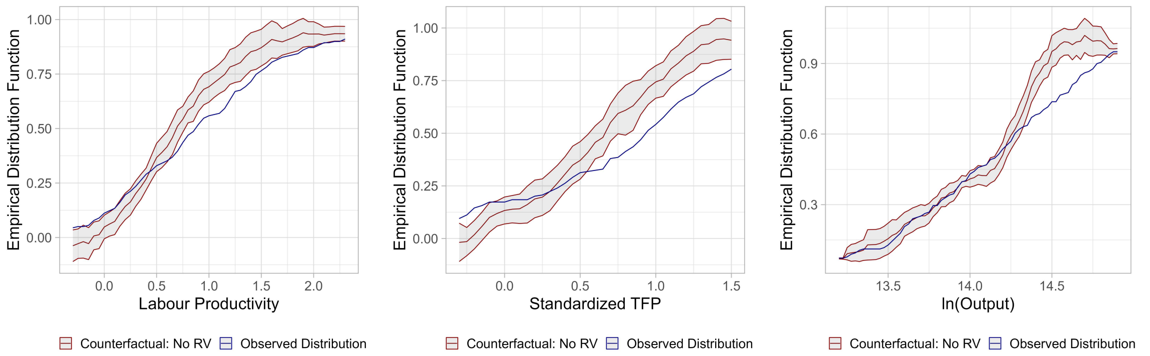


Figure 4.14: Effect on Mine Survival

Note: Based on equation (4.9) and a balanced mine-year panel. For blue line: Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

First, I use firm-level financial data to test whether financial constraints are lifted by the policy. Figure 4.16 presents the following for a sample of all stock-market listed firms (blue) and firms that have coal production as only business (red): Firms with a higher exposure to the policy, i.e., a higher share of closed capacity, experience a larger drop in their debt ratio (debt over the sum of debt and equity) by up to 15 p.p. (see Panel A). The reduction in financial constraints also simultaneously translates to higher stock values and dividends (Panel B and C), i.e., financial markets expect firms' new financial potential to restructure their production to result in improving firm result metrics.

I further provide support on the financial constraint mechanism in Table 4.2, where I explicitly test the role of financial constraints for the extent of reallocation due to the policy. Panel A gives the baseline results from Section 4.5, i.e., surviving mines of policy-uptaking firms increase output, employment, and capital. In Panel B, I add an additional difference-in-differences interaction, that gives the policy exposure measured in the net premium per remaining tonne of production ($[Net\ Exposure\ in\ DM_j] \times 1[Year > 1965_t]$) instead of the output-weighted measure, $[RV\ Exposure_j]$. The net premium accounts for the heterogeneous payments to competitors for their closures and also considers different closure subsidies depending on whether the closure belonged to the early part of the policy (*Vorausaktion*) or not. The firms of 43% of the mines in the sample are net receivers from the policy. Hence, here I run a horse race of the exposure to the policy measured in the share of closed quantity, $[RV\ Exposure_j]$, versus the actual financial exposure, $[Net\ Exposure\ in\ DM_j]$. It can be seen that only the monetary exposure matters for the reallocation effects. Hence, the net amount of subsidies received is driving the reallocation process. The more money earned per remaining unit of production, the stronger the reallocation and output increase.



(a) Labour Productivity (b) TFP (c) Output

Figure 4.15: Distributional Effect of *RV* Policy

Note: Based on distribution regression approach by Chernozhukov et al. (2013). Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

In Panel C, I explicitly test whether the lifting of financial constraints matters for the policy response. The policy-induced increase in input usage is stronger for those mines that belong to firms with a higher debt ratio (debt capital divided by overall capital), i.e., that were more financially constrained, in 1961 right before the policy.

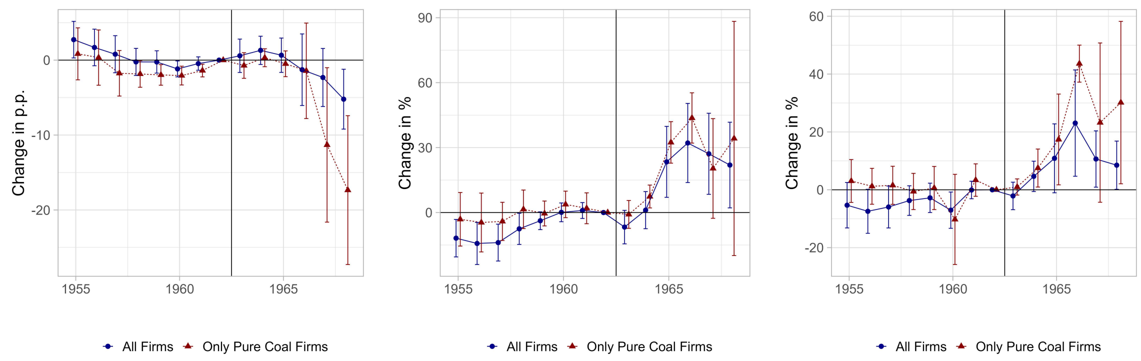
In Panel D, I show that the effect is not driven by the fact that output spillovers just reflect spatial spillovers between mines of the same firm or across firms. The existence of closed mines of other firms nearby has no explanatory power across firms, only the exposure of the mines' own firm matters.

Investments, Infrastructure, and Technology Adoption. I subsequently show how the lifting of financial constraints mapped into productivity gains by leading to more investments, an improved mine infrastructure, and a higher technology adoption. In the context of mines, productivity increases require investments, e.g., new tunnels to rich coal layers. While such construction work is very costly, firms could use the earned closure premium to invest.

The left panel of Figure 4.17 documents that the remaining mines of uptaking firms on average slightly increase their mine depth by 4%. This effect is especially driven by very deep mines. The probability that mines have a maximum depth of over 1,050m (75th percentile) is increasing by 10 p.p. after the policy. On average these deep mines are younger and more productive, so that investments might have a higher return. Deepening a mine is very costly, can take several years, and can be seen as a large-scale investment into the infrastructure of a mine.

Moreover, the number of conveyor tunnels, i.e., vertical tunnels through which coal is brought to the surface, is unaffected. This means that current mines are extended within the already existing mine framework (e.g., increasing depth) instead of expanding the mine across its former borders or acquiring new, deep coal fields. The latter would most likely require new conveyor tunnels.

In line with this evidence on investments into the mine, the middle panel of Figure 4.17 shows

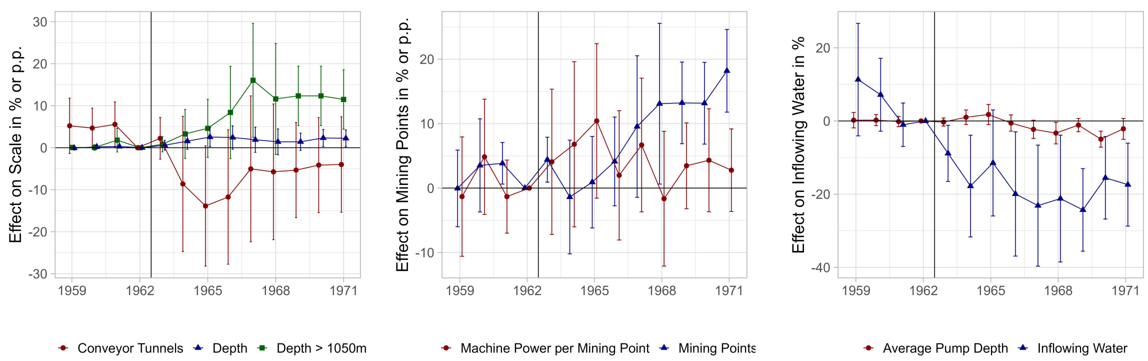


(a) Debt Ratio (b) Stock Values (c) Dividends

Figure 4.16: Effect on Debt Ratio and Stock Market Evolution

Note: This plot documents how the policy uptake affects firm-level stock values, dividends, and the firms' debt ratios. Right panel based on equation (4.9) at the firm level. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported. Sample ends in 1967 as the majority of firms merged to the agglomerate Ruhrkohle AG afterwards.

that there is an increase in the number of mining points, i.e., the positions in the mine at which coal is quarried at the same time. Hence, within the mine, work is done at more locations. Setting up a mining point requires investments in its setup. Even though the number of mining points increases, the machinery power per mining point does not decrease. Hence, the infrastructure investments into the mine go hand in hand with more, well-equipped mining points.



(a) Infrastructure Investments (b) Capital Investments (c) Water Management

Figure 4.17: Effect on Measures of Investments

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

The right panel shows that the amount of water flowing into the mine is reduced after the policy. As pit water is an important security threat, mines improve water management and worker safety. Also, the average pump depth falls. Mines pump water from nearer to the surface reducing the risk of flooding in deeper mine parts. In the Appendix, I provide fur-

Table 4.2: Mechanism

	log(Output) (1)	log(Miners) (2)	log(Machinery) (3)
Panel A: Baseline			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.318^{***}	0.299^{**}	0.376^{***}
	(0.100)	(0.111)	(0.139)
Panel B: With Net Exposure			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.091	0.010	0.121
	(0.128)	(0.151)	(0.124)
$[Net\ Exposure\ in\ DM_j] \times 1[Year > 1965_t]$	0.005^{***}	0.006^{***}	0.005^{**}
	(0.001)	(0.002)	(0.002)
Panel C: Financial Constraints			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	-0.215	-2.227^{*}	-2.084^{*}
	(1.042)	(1.185)	(1.013)
$[RV\ Exposure_j] \times 1[Year > 1965_t] \times [Debt\ Ratio_j]$	0.834	4.004^{**}	3.628^{**}
	(1.643)	(1.926)	(1.613)
Panel D: Local vs. Within-Firm Spillovers			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.303^{***}	0.283^{***}	0.387^{***}
	(0.098)	(0.103)	(0.123)
$1[Closure(s)\ Other\ Firm\ [0, 2.5)\ km_j] \times 1[Year > 1965_t]$	-0.030	0.070	0.064
	(0.061)	(0.065)	(0.097)
$1[Closure(s)\ Other\ Firm\ [2.5, 5]\ km_j] \times 1[Year > 1965_t]$	-0.065	-0.039	0.066
	(0.044)	(0.051)	(0.064)
Mine FE	Yes	Yes	Yes
Coal District-Year FE	Yes	Yes	Yes
Observations (Panel A, B & D)	1,012	1,012	861
Observations (Panel C)	633	633	534

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Significance levels of 10%, 5% and 1% are denoted by *, ** and ***.

ther evidence of improved worker safety as the prevalence of accidents slightly decreases with policy exposure (see Figure 4.A.15).

Further, given the investments and increased machinery power, I investigate changes with respect to technological change. In the 1950s to 1970s, there was a major switch from manual, non-mechanized coal mining (i.e., workers with automatic hammers) to large-scale machinery usage. While on average 33% of output was produced by mechanized production methods in 1957, this share increased to 90% in 1971. In Figure 4.18, I show that mines with a high policy exposure increase the share of mechanized production after the policy while the share of mechanized production points is unaffected. Hence, the policy led to an increasing use of new technologies such as cutting and pealing machines.

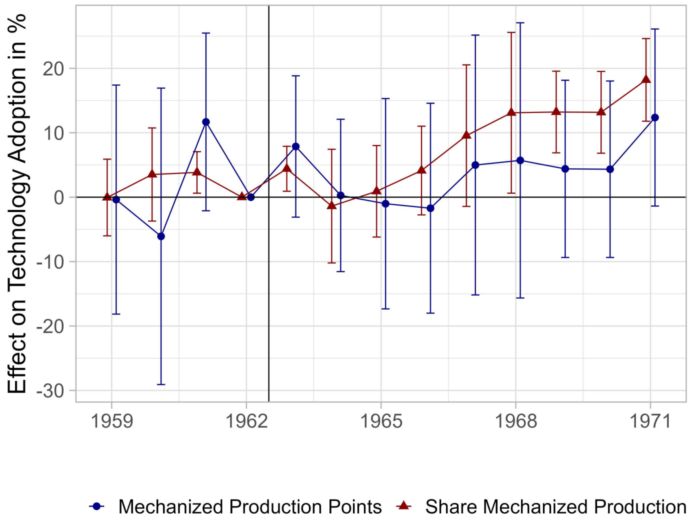


Figure 4.18: Effect on Mechanization

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

4.5.5 Spillovers to Downstream Cokery Industry

The production reallocation also has relevant effects on the downstream industry. I here focus on the important downstream industry that refines coal to coke, an intermediate good used in steel production. About 40% of the coal quarried is used for coke.

I study whether mines react to the policy by changing the type of coal they quarry. The type of coal is relevant as, for example, only some types of coal can be used for further refinement to downstream products such as coke. German cokeries became the most important demander for German coal in the 1960s and 1970s and offered stable demand given that coke could not be substituted with oil. Hence, I study whether mines reallocate to producing fat coal, which is the coal type primarily used in cokeries. Figure 4.A.16 shows that there is an increase in fat coal output by about 10% while non-fat coal output is unaffected. Thus, treated mines are able to transform their production towards coal varieties which are more stable and less prone to demand fluctuations. Hence, the policy allows firms to reallocate production towards more stable markets.

In line with the spillovers to mines with cokeries and the increasing output of coal that is useable for coke production, Figure 4.A.17 in the Appendix shows that exposed mines increased employment in their vertically integrated cokeries. Hence, reallocation has an impact on the employment decisions in the downstream industry, too. A back-of-the-envelope calculation suggests that treated mines increased their cokery employment by in total 350-400 jobs until the end of the 1960s.

At the same time, the right panel of Figure 4.A.17 provides no evidence of an improvement in cokery input quality and cokery efficiency. I proxy input quality with the share of volatile content in the coal used for coke production. The higher this share, the more energy is lost in the cokery. Cokery efficiency is measured as the output-input ratio of coke relative to coal, i.e., how much coke is produced from a unit of coal.

4.6 Other Potential Mechanisms

In this section, I discuss other potential mechanisms and their role.

Bargaining Power within Syndicate. Given the existence of joint retail organizations in this industry, one concern is that positive output spillovers to mines within the firm are a mechanical outcome of (legal) negotiations between firms. However, gained bargaining power from mine closures, which result in output increases, should lead to fewer shifts cancelled due to insufficient demand for treated firms relative to non-uptakers. Contrary to this, Figure 4.A.18 in the Appendix provides no evidence for this. Neither the extensive margin of cancelled shifts due to insufficient demand nor the intensive margin differ from mines of other exposure levels.

Local Demand and Long-run Contracts. First, a firm-specific or local demand shock could explain production increases in the remaining mines of exiting firms after the policy. However, all mines are located close by and likely are exposed to the same shocks. Also, joint retailing through the retail organizations has, by law, the purpose and mandate to smooth regional and coal type-specific fluctuations in demand across mines (High Authority, 1956*a*). Second, as the *RV* policy was unforeseen, one concern is that coal firms still have running long-run contracts to fulfill. Increasing output levels in the remaining mines could be one reason then to reach the necessary output levels. However, coal firms did not independently sign supply contracts with coal-demanding entities. Instead, they only sold coal through the retail organizations (High Authority, 1965), making this concern obsolete.

Political Influence and Workers' Bargaining Power. Differences in regional politics and mine-level bargaining power of workers could drive my results. Political parties had different opinions on policymaking in the coal industry. While the governing Christian democrats and Liberals passed the *RV*, the only opposition party, the Social Democrats, was in favour of less severe, short-run employment drops (Bundestag, 1959*b*, 1962). Therefore, I analyse whether local heterogeneity in political attitude affected firms' reallocation decisions - e.g., towards fewer layoffs. Using pre-policy voting results in the 1961 federal election at the county level, Figure 4.A.19 in the Appendix shows that there is no robust effect of political attitude on the reallocation decision.

To test whether spillovers are determined by workers' bargaining power, I exploit mine-level heterogeneity in workers' elasticity of labour supply. I use mine-level data on the share of foreign workers (so-called *Gastarbeiter*) which varied between 0 and 36% between mines. Foreign workers were less likely to leave the industry relative to native workers and it was difficult to organise them in unions (Seidel, 2014). Figure 4.A.20 in the Appendix shows that there is no difference in the effect on inputs and outputs between mines with different *Gastarbeiter* exposure before the policy.

Output Quality. A different explanation for the diverging developments of mines over time is a change in the quality of products over time. Admittedly, coal is very homogeneous but perceived quality can vary. As this is unobservable or hard to measure, quality usually is proxied by input price data (De Loecker et al., 2016) or measures based on demand assumptions (Amiti and Khandelwal, 2013, Khandelwal et al., 2013). As input prices, which vary at the mine level and over time, I only observe wages. Wages, however, can for example be distorted from bargaining differences.
Instead, I use a measure that is specific to my setting, the amount of non-coal output quarried. The more non-coal output was quarried, the more likely it was that sold coal included non-coal content even after the washing process. By that, the energy content of coal supply is affected. Figure 4.A.21 in the Appendix shows that the amount of non-coal content is unaffected by the treatment.

4.7 Robustness Checks

Subsequently, I provide a number of empirical robustness checks.

Long-run Effects. In my main analysis, I restrict my sample to years up to 1971 because of other policy changes in the industry afterwards. For an extended sample until 1980, I show that the effects on output and input decisions are persistent (see Figure 4.A.22).

Production Function Approach. My TFP results rely on the correct specification of the production function. In Section 4.5, I provided results based on production function variants with either pit wood or electricity as the material input. In Table 4.B.7, I provide further robustness checks. Next to my baseline results in columns (1) and (2), I estimate a translog production function with pit wood as material in column (3) to flexibly account for potential substitution patterns between wood, labour and capital. In columns (4) and (5), I extend the time horizon beyond 1971 to 1980 and estimate the effect on TFP and labour productivity. In column (6), I take account of a potentially changing production function over time by adding two interactions of a time trend with the labour and capital variable respectively to the Leontief production function. This should also take care of potentially factor-biased productivity gains (De Loecker et al., 2020). Lastly, in column (7), I explicitly include a measure of technology adoption, i.e., the share of mechanized production at the mine-year level, as an input in the production function to account for factor-biased technology change. All of these specifications support that the policy led to productivity gains. Similarly, Table 4.B.8 shows that the extensive margin results, i.e., that productivity drives exit, are robust to the different TFP estimates.

Markdowns. Higher markups could also stem from endogenous input prices, i.e., for example, lower wages as workers face a reduced labour demand (Avignon and Guigue, 2023, Morlacco, 2019, Rubens, 2023*b*). However, I show in Figure 4.A.23 of the Appendix that average wages increase in response to the policy in highly- relative to less-treated mines. While this does not eliminate the possibility of rising markdowns due to potentially an increasing

marginal revenue product of labour in treated mines, this evidence limits the concern.
Additionally, I calculate markups based on the output elasticity from the Cobb-Douglas production function with electricity as material input. In electricity markets, market power is limited, i.e., one may assume identical input prices across mines. In Figure 4.A.24 in the Appendix, I show that the pattern of rising markups due to decreasing marginal costs is also evident in this robustness check - also if I restrict the sample to mines with their own power plants (self-suppliers). By construction, changes in markups over time should not stem from heterogeneous markdown developments for treated and non-treated self-suppliers.

Staggered Exit and Difference-in-Differences with Continuous Treatment. While all firms had to decide whether to close mines or not until late 1964, exit happened in a staggered fashion with most exits between 1964 and 1966 (see Figure 4.5 above). To account for this, I rerun my main analysis in a staggered difference-in-differences event study. In Figure 4.A.25 in the Appendix, I show that my results on the quantity spillovers as well as input usage are confirmed by the staggered adoption model.
Note that recent literature (Borusyak et al., 2024, Callaway and Sant'Anna, 2021, Callaway et al., 2024*b*, De Chaisemartin and d'Haultfoeuille, 2020, Sun and Abraham, 2021) has shown that event study results (with continuous exposure) can yield distorted estimates of the treatment effect. However, the relatively high share of never-treated units in my analysis reduces this concern (Borusyak et al., 2024).
Further, Callaway et al. (2024*a*) show that the difference-in-difference model with continuous exposure only identifies a causal effect under a more demanding parallel trends-type assumption. Therefore, I repeat my main analysis in a standard binary difference-in-differences model where I compare mines with strictly positive exposure ($[RV\ Exposure_j] > 0$) to mines with zero exposure (see Figure 4.A.26 in the Appendix). Results are qualitatively identical.

Non-Linearity in Treatment Effects. My regression design implicitly assumes that the marginal effect of an increase in the treatment exposure is constant independent of the level of the treatment. To ensure that this does not blur the estimated results, I (i) estimate separate treatment effects for mines with below/above median exposure among the treated mines and (ii) test for quadratic relationships between treatment and outcomes in the Appendix (see Table 4.B.6). Both tests indicate that the spillovers are especially driven by highly treated mines.

Exit, Sample Composition, and Selection into Treatment. To capture the whole industry, my analysis included all mines, which were not closed through the policy, in the regressions. However, some mines have been closed before or after the policy and two mines opened in the 1960s, so that the composition of the control and treatment groups varies over time. To ensure that this selection process does not affect my estimation results, I rerun my main analyses in the Appendix for a balanced panel of mines that operated throughout the whole sample period (see Figure 4.A.27). I further reproduce my main analysis for a subsample of only multi-mine firms. This ensures that selection into treatment, which is only possible for multi-mine firms as another mine of the same firms needs to be shut down, is not driving the results (see Figure 4.A.28).

Identification, Inference, and Weighting by Size. Up to now, I identified spillovers by comparing mines of firms having different treatment exposures within coal districts ($N = 3$) using coal-district fixed effects. In Table 4.B.9 in the Appendix, I show that my results are not sensitive to comparing all mines (only year fixed effects) or within more refined coal areas ($N = 7$) which split the large *Ruhr* district into subregions (coal region-year fixed effects). That the results are unaffected by the regional identification cell is further support for the absence of spatial spillovers.

With respect to inference, my results are unaffected by using standard errors clustered at the mine level, at the pre-treatment owner level, or using spatial standard errors (Conley, 1999) instead.

Lastly, I rerun my main estimations of the within-firm analysis for regressions weighted by size, i.e., output in the pre-policy year 1962. This is motivated by the higher relevance of changes in larger mines. Table 4.B.10 in the Appendix shows that the treatment effects are slightly larger (however, not always statistically significantly larger) in the weighted regressions. This is also in line with the reallocation towards larger mines.

4.8 Discussion and Conclusion

In this section, I assess the performance of the closure subsidy relative to other alternative policies and discuss the implications of the paper. Different industrial policies can vary substantially in their economic effects.

Net Employment Effects. Beyond productivity, a policymaker might also care about welfare effects such as those on the labour market which I only marginally considered up to now. At first glance, there is a trade-off between the productivity-oriented policy (i.e., fewer input usage and mine closures of unproductive mines) and employment in the short run. However, this might not be the case in the long run as productivity gains can lead to mines surviving for a longer time and jobs being saved. To incorporate this, I conduct a back-of-the-envelope calculation of net employment effects due to the policy.

The policy led to the closures of mines which employed 87,000 workers right before the policy in 1961. My spillover analysis suggests that actually only 33% of these jobs got lost (i.e., 29,000 jobs).[16] The rest is recovered in the remaining mines. This also already accounts for job loss through the change from labour- to capital-intensive production due to technology adoption.

Further, the spillovers led to productivity gains, so that remaining mines on average survived for six more years. Given that treated mines on average survive 26 years post-policy and make up about 50% of the industry production at the end of my panel in 1971, long-run job savings are substantial. Treated mines employed between 85,000 and 126,000 employees per

[16]Note that this estimate is based on the partial equilibrium assumption that firms that are treated with zero exposure did not reduce their output in response to the policy (e.g., due to lower productivity relative to treated mines). However, my spatial spillover analysis supports this assumption.

annum throughout the post-policy years 1963-1971. Smoothing, e.g., 85,000 saved jobs in six years over 26 years, translates to a conservative estimate of 20,000 jobs per annum. This almost fully compensates the job loss due to the policy (29,000 jobs).

Note that job loss due to the decline of the industry would have occurred anyway in closed mines. For example, mines of the upfront part of the policy, *Vorausaktion*, closed without knowing of the premium and earned it ex-post. Thus, the almost full compensation is a lower bound, conservative estimate for the net job gains from the policy. Hence, the policy was not detrimental to employment in the aggregate. However, there are important distributional implications over time with early mass layoffs and late savings.

To also account for across-industry employment spillovers, I provide Figure 4.A.5. There are no significant spillovers to other industries in the studied time period.

Closure Subsidy vs. Wage Subsidy. As an alternative to a closure subsidy, the government could sustain jobs by directly subsidizing wages.[17] Instead of downsizing the industry, the government could try to sustain the industry. I calculate back-of-the-envelope costs of a non-discriminating wage subsidy for all mines and a discriminating version for those mines with high exit probabilities only.

Our extensive margin IV regression from Table 4.1 shows that a one standard deviation increase in labour productivity (mean (sd): 305 (52) tonnes per worker per annum) causes a decrease of 37 p.p. in the probability of exit. Non-discriminatorily subsidizing every sixth shift then implies an effective increase in labour productivity by one standard deviation. In the best scenario, this could lead to not a single mine exiting the market at the time of the policy (31% of mines exited through the policy). Given the fiscal closure premium budget of 350 mio. DM, an average wage of 25.11 DM per shift in 1961 and about 98 million shifts in 1961, this however would take more than the overall premium budget from the closure policy for wage subsidies of just one year. Similarly, subsidizing wages at the employment level as of right after the closure policy given the policy-induced market exit would have been too costly.

In a world where the policymaker knows who will exit, it could target subsidies to those mines. Paying every sixth shift for only those mines that exit through the RV would allow the policymaker to pay a subsidy for about three years.

Hence, pure wage subsidies - even if they are targeted - cannot persistently save jobs in my setting.

Buy Excess Coal or Subsidize Coal Prices? The policymaker could also increase its own coal demand to save the industry. However, this is too costly. Just the excess, not-sold coal that was stored on pithead stocks between 1964 and 1966 made up 13.6 million tonnes. At an average price of around 80 DM/tonne in these years, purchasing the excess coal for just three years would have cost three times the policy budget. Further, this did not even consider that excess coal production would likely have been higher absent the exit policy.

[17] This is, for example, currently proposed as policy for the declining lignite industry in East Germany (German Federal Ministry for Economic Affairs and Climate Action, 2019).

Similarly, an alternative policy could have been to subsidize the price. In fact, this policy was also debated in the parliament (Bundestag, 1965). Assuming the most favorable condition, i.e., full pass-through of price subsidies to consumers, the excess demand of 13.6 million tonnes would require a 2% price cut (following my demand elasticity estimates in Table 4.B.1). A 2% price subsidy throughout the first three policy years would, however, cost 670 million DM, i.e., almost twice the government payments for the closure transfers.[18] In fact, Storchmann (2005) shows that policy interventions in the industry after 1970 (mostly price subsidies) were much more costly than the closing subsidy. Again, the volume of excess coal would likely have been higher without the policy.

Economies of Scales and Mergers? Given that large mines usually are more productive, mergers could increase productivity, too, but might be less costly for the policymaker. Throughout my sample, more than twenty mine mergers took place. Since mergers in this industry require that mines are geographically located next to each other, these mine mergers mainly took place within firm among mines of similar productivity level.[19] In Table 4.B.11 in the Appendix, I show that they barely affect mine-level outcomes. I only find evidence for a reduction in employment but no effect on productivity, output, capital stock, and survival. Hence, this type of merger would not improve mines' productivity and also lead to employment drops.

Entry? The industry could also gain productivity by opening new mines that dig in very profitable coal fields. The high fixed costs of setting up a mine (several years of preparation) could be financed by the government if firms themselves do not want to enter. However, mines were already quarrying coal in the most northern part of the Ruhr area, where coal layers were the thickest and most yielding in Germany. Hence, productivity improvements by opening new coal fields with better geological preconditions would not have been possible.

Policy Improvements. For the future implementation of similar policies in other settings, it is crucial to understand which policy details could have been improved. First, one result of the *RV* is increasing productivity dispersion among surviving mines, i.e., weaker mines remain in the industry, too. The policy could have targeted the exit of such inefficient mines by, for example, introducing heterogeneity in the subsidy by mine or firm size - as size highly correlates with productivity. Further, the policy could have steered the extent of exit by changing the average subsidy size. Lastly, the policy was half financed by competitors paying the exit subsidy. However, quantity spillovers to competitors were limited as I showed above. This raises the question whether the policy should have had a smaller premium participation by competitors. Competitors only profited from the industry-level reduction in overcapacities smoothed across firms through the retail organisations.

Further, I showed that the policy triggered reallocation towards already large and productive mines and the productivity gains centered in these entities. Thus, large, productive mines

[18] In fact, the German government subsidized coke coal sales in the 1970s and paid almost 1 billion DM for it.

[19] Also, across-firm acquisitions were hardly possible given the break-up of the industry after World War II (Allied Higher Commission, 2019).

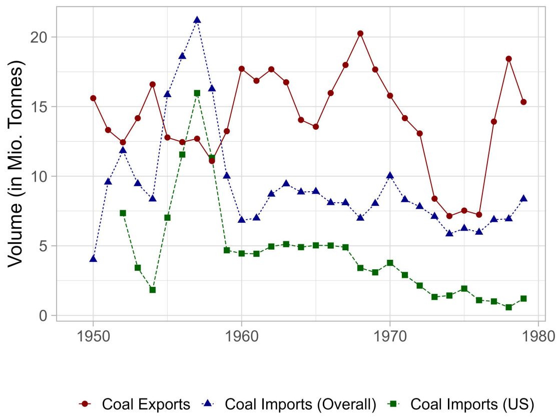
formed endogenously. A more precise focus on not incentivizing the exit of mines of this type could have been a more targeted policy.

Conclusion. In this paper, I analyse how industrial policy steers exit at the example of an economization scheme in the German coal mining industry. I find that the policy let the 'losers' go and endogenously raised the performance and market share of 'winners' in the industry. The policy fostered technology adoption and productivity gains along various margins: exit of inefficient mines, within-mine productivity gains and reallocation towards large, productive mines. The policy's costs are compensated by marginal costs savings in the industry. More productive mines survive longer prolonging the industry's lifespan and saving jobs in the long run.

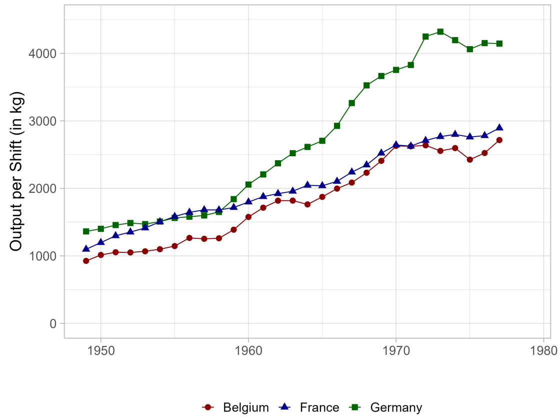
This evidence motivates the consideration of exit subsidies as one way to persistently improve an industry's productivity. In contrast to common price or wage subsidies, which often are meant to especially help struggling firms in an industry to keep them alive, this type of policy instead promotes and selects productive firms. My findings are relevant to many industries that are currently in decline such as steel production, non-renewable energy production, or car manufacturing, where optimal mechanisms for capacity reduction are a common debate. For example, in Germany, coal power plants are paid for market exit given the country's goal to a green transmission. In some of these industries, policy-driven productivity gains might be sufficient to keep them alive (for longer), in contrast to the coal industry at hand.

My results, further, are of interest for non-declining industries with temporary overcapacities such as milk, fishing, wine, or vegetable production where incentives for market exit are common policy tools (e.g., Commission of European Union, 1988, Council of European Union, 2006, Raggi et al., 2015, Commission of European Union, 2016). Lastly, my findings have insights for industries in which the policymaker might want to decrease overall production (e.g., phase-out of non-renewable energy production) with increasing productivity and efficiency of the remaining firms at the same time.

Appendix A: Additional Figures

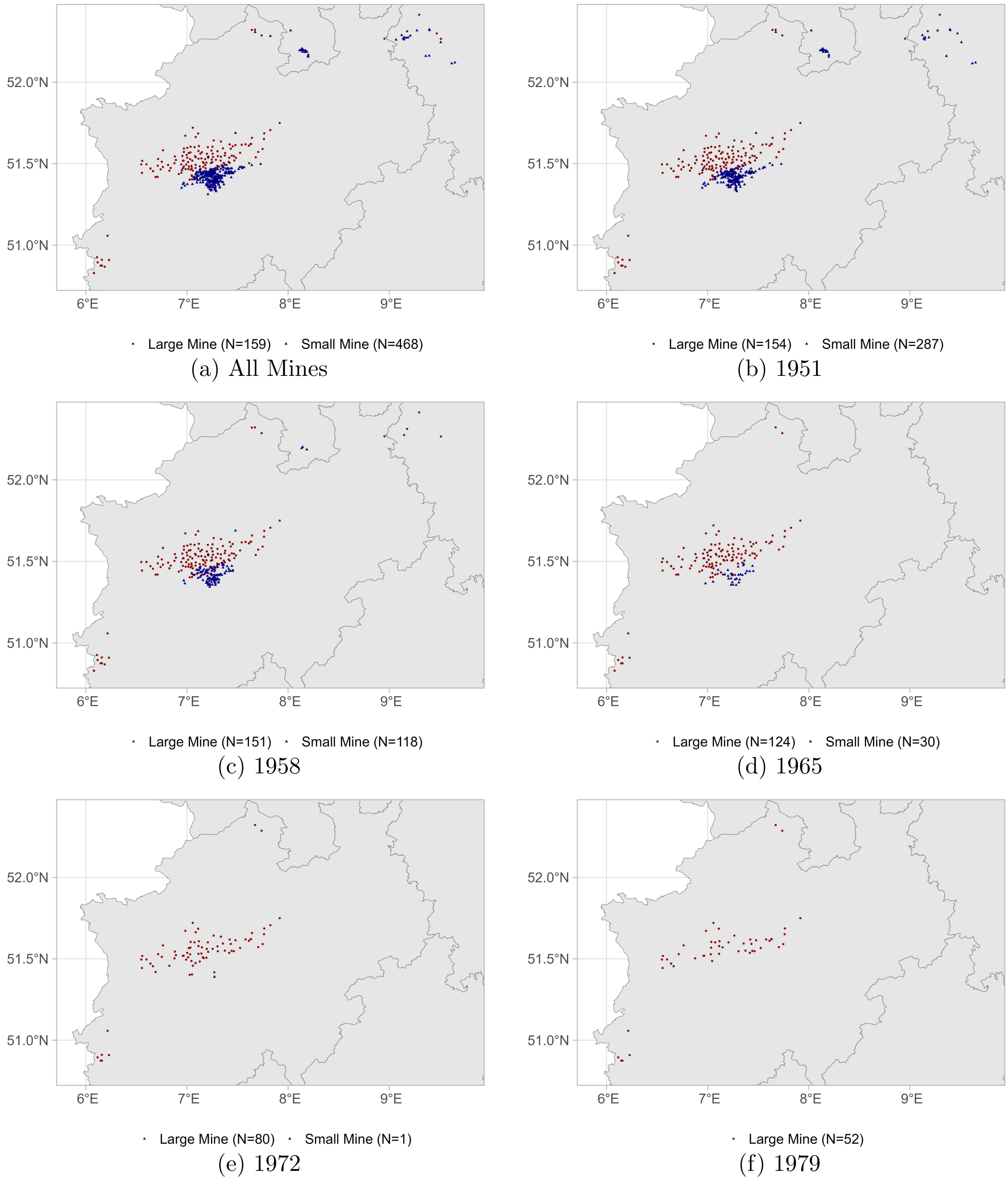


(a) Import/Export Volume



(b) Labour Productivity Across Countries

Figure 4.A.1: Relevance of Cross-National Coal Trade



(a) All Mines

(b) 1951

(c) 1958

(d) 1965

(e) 1972

(f) 1979

Figure 4.A.2: Mines over Time

Note: These plots show the location of active mines by year. Mines are classified into large and small mines. Large mines include all mines for which detailed production data is available. Small mines are all other mines. All mines included which operated at least for one year after 1951. For merged mines, I count the joined mines separately in their original independence.

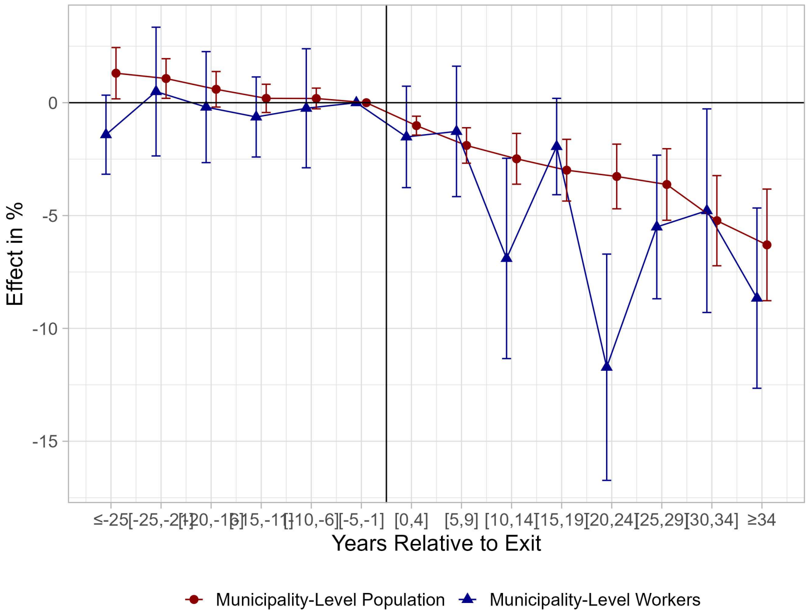
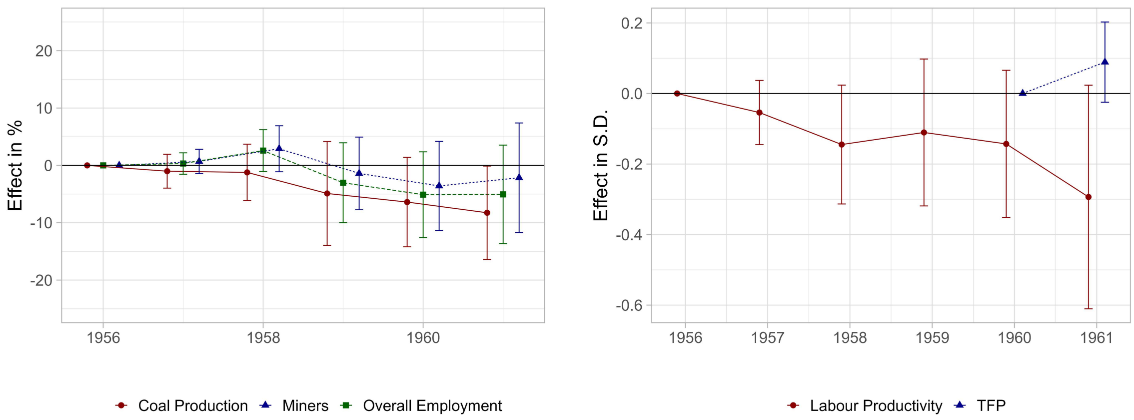


Figure 4.A.3: Effect of Mine Closures on Municipality-Level Population

Note: This plot documents the effect of a mine closures on the municipality-level population and workers. The sample is restricted to municipalities in the state of Northrhine-Westphalia, where all coal mines are located (except for few in Lower Saxony). I harmonize municipality boundaries to borders as of today. I weight observations by today's (2022) municipality population. The regression includes municipality fixed effects as well as year fixed effects. For data on working population ('Erwerbstätige'), only census data is used. Standard errors are clustered at the municipality level ($N = 396$) and 90% confidence intervals are reported. The treatment is the number of mines per inhabitant in the year of the mine closure. Hence, the regression has the format:

$$Y_{mt} = \alpha_m + \gamma_t + \sum_{\tau=-6,\tau\neq-1}^{8} \beta_\tau [\frac{\#Mines\ Closed_{m,t-\tau}}{Population_{m,t-\tau}}]_{m,t-\tau} + \epsilon_{mt} \tag{4.11}$$

where m and t give an index for municipality and year. Municipality and year fixed effects are given by α_m and γ_t. Endpoints are binned. The index of the leads and lags used in the event study is τ. Mine closures are closures of those mines for which productivity data is available, i.e., large mines. The coefficients β_τ of the regressions are multiplied with the median value of $med\{\frac{\#Mines\ Closed_{m,t}}{Population_{m,t}} | \frac{\#Mines\ Closed_{m,t}}{Population_{m,t}} > 0\}$, so that the elasticities in the event study can be interpreted as the effect of one additional mine closure.

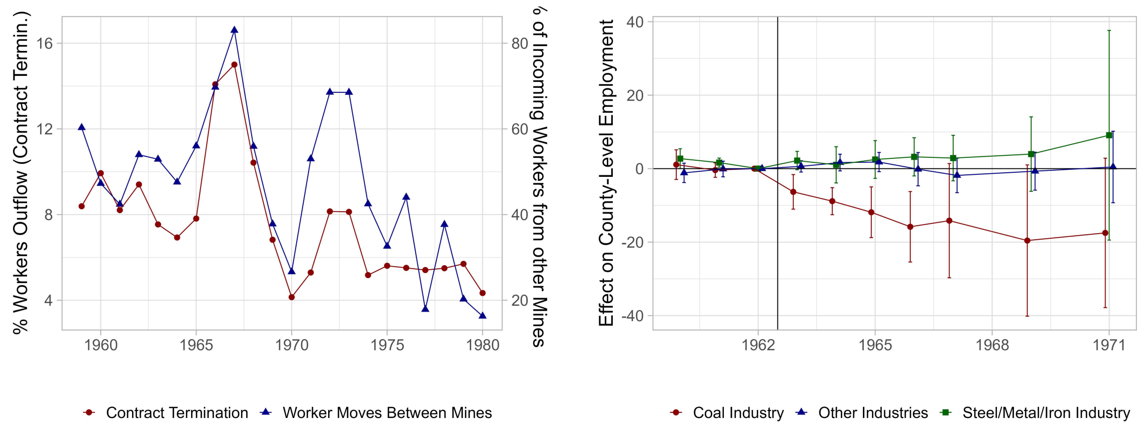


(a) Output/Employment

(b) Productivity

Figure 4.A.4: Anticipatory Effects

Note: These plots document how mines, that were closed through this policy, developed in the years prior to the policy (until 1961). I run simple regressions of logged output, employment and standardized productivity measures on mine and district-year fixed effects as well as an dummy for policy-uptaking mines interacted with year fixed effects. No observations of mines in the year of closure are included. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.



(a) Worker Turnover

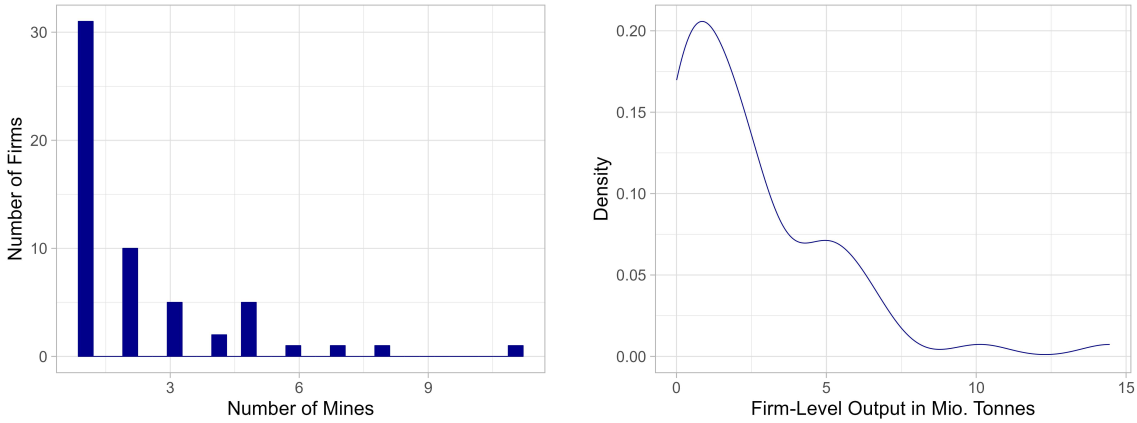
(b) Effect of *RV* on County-Level Employment

Figure 4.A.5: Worker Turnover

Note: The left plot documents two time series: First, the share of workers whose contracts have been terminated (either dismissed or voluntary leave) among all workers per year. Second, the share of workers among all new incoming workers who have been working at a mine before. The right plot documents county-level estimation results of a difference-in-differences estimations of the format:

$$Y_{ct} = \alpha_c + \gamma_{rt} + \sum_{t=1960,\neq 1962}^{1971} \beta_t [\frac{\#Mines\ Closed\ RV_c}{Population_{c,1962}}]_c \times 1[Year = t]_t + \epsilon_{ct} \tag{4.12}$$

where c and t give an index for county and year. Municipality and Regierungsbezirk-year fixed effects are given by α_c and γ_{rt}. The coefficients β_t of the regressions are multiplied with the median value of $med\{\frac{\#Mines\ Closed\ RV_c}{Population_{c,1962}} | \frac{\#Mines\ Closed\ RV_c}{Population_{c,1962}} > 0\}$, so that the elasticities in the event study can be interpreted as the effect of one additional mine closure. Standard errors are clustered at the county level (counties as of 1971) and 90% confidence bands are reported.

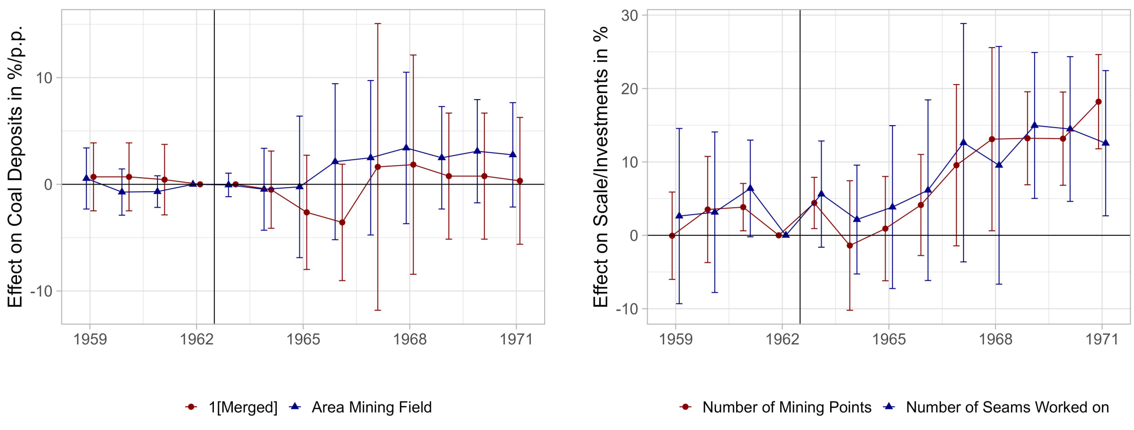


(a) Number of Mines per Firm

(b) Output per Firm

Figure 4.A.6: Market Structure

Note: These plots document the distribution of firm size in the industry in the pre-policy year 1962 (small mines, 'Kleinzechen' not included).

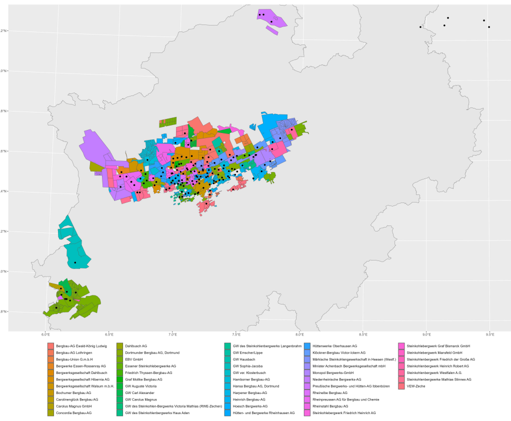


(a) Tests of Mine Expansion

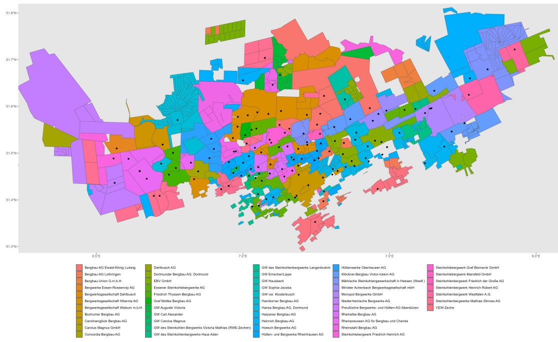
(b) Within-Mine Capacity Changes

Figure 4.A.7: Effect on Capacity Proxies

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. This plot documents the effect of the policy uptake on the coal field size and the number of seams of the remaining mines of the same firm. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.



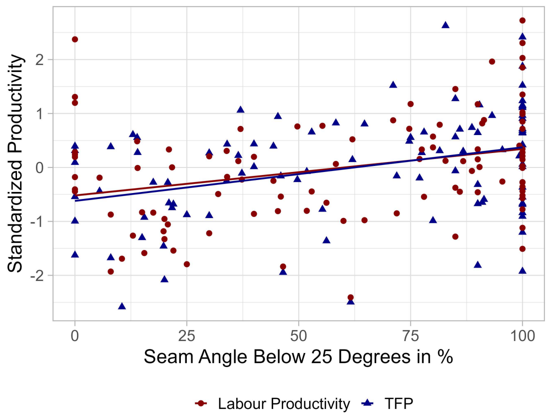
(a) Entire Sample



(b) Ruhr Area Only

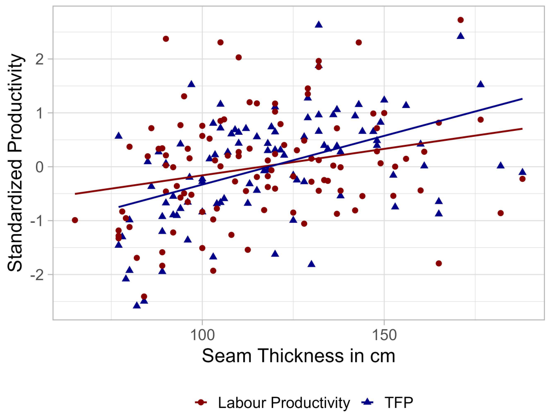
Figure 4.A.8: Coal Field Ownership in 1962

Note: These plots document the 'Berechtsame' (i.e., coal field ownership) by firms before the policy in 1962. The administrative borders of Northrhine-Westphalia are included.

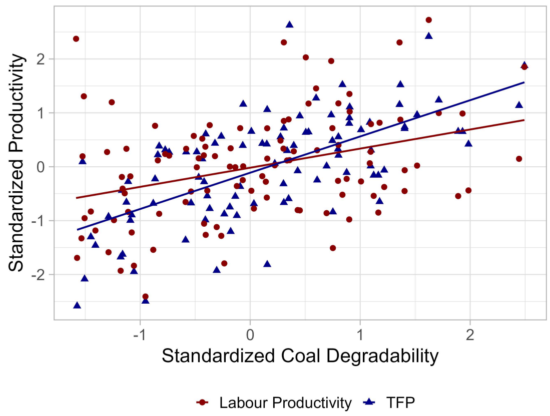


(a) Coal Angle

(b) Marl Thickness



(c) Seam Thickness



(d) Coal Degradability (PCA)

Figure 4.A.9: Correlation of IVs and Productivity

Note: These plots give the correlation of the IVs with productivity measures separately and jointly (from predicted values of a principal component analysis). Labour productivity is calculated as the average across the policy-premium relevant years 1959-1961.

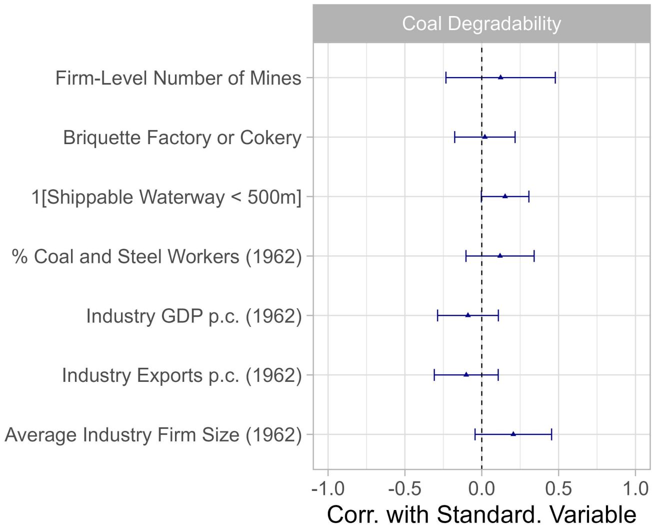


Figure 4.A.10: Correlation of IV and Mine-Level or Regional Characteristics

Note: This plot documents the correlation of the mine-level IV with mine-level or regional information for the years 1961 or 1962. Data on industry GDP and exports per capita as well as the share of industrial workers comes from (Statistisches Landesamt Nordrhein-Westfalen, 1964). Distance to nearest shippable waterway is calculated based on the shapefile provided by (Wasserstraßen- und Schifffahrtsverwaltung des Bundes, 2021).

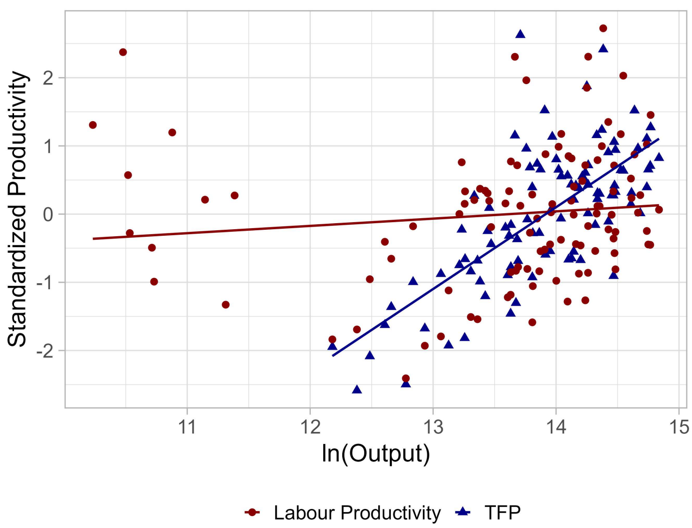


Figure 4.A.11: Correlation Firm Output and Mine Labour Productivity

Note: This plot documents th correlation between firm-level output and mine-level labour productivity based on 1961 data for all mines which have been operating for the full year.

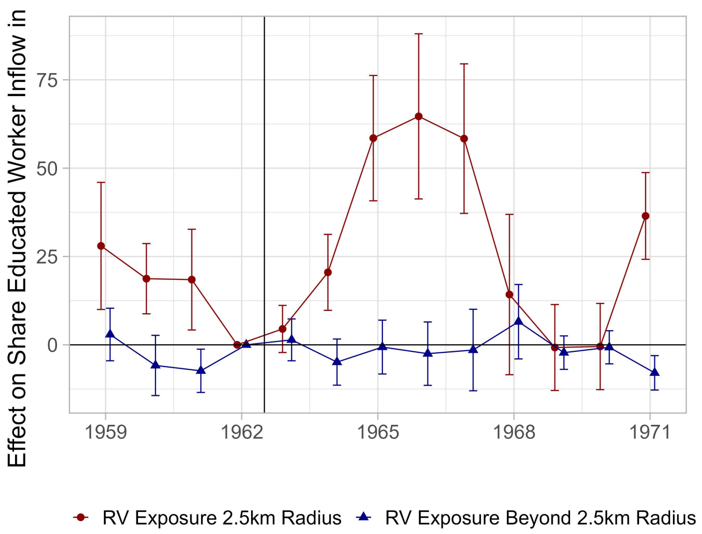
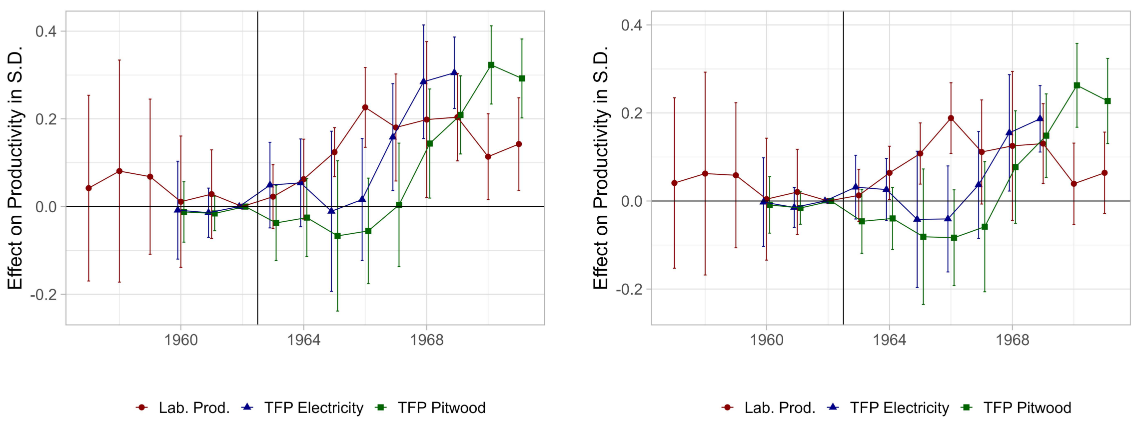


Figure 4.A.12: Effect of RV on Share of Educated Workers Among Joining Workers

Note: Based on equation (4.9) where the treatment exposure is splitted in $[RV\ Exposure_j]$ below and up 2.5km around a mine. Coefficients multiplied with mean $[RV\ Exposure_j]$ within 2.5km for blue line and beyond 2.5km for red line for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.



(a) Without Output

(b) With Output

Figure 4.A.13: Effect on Productivity - With and Without Controlling for Output

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.

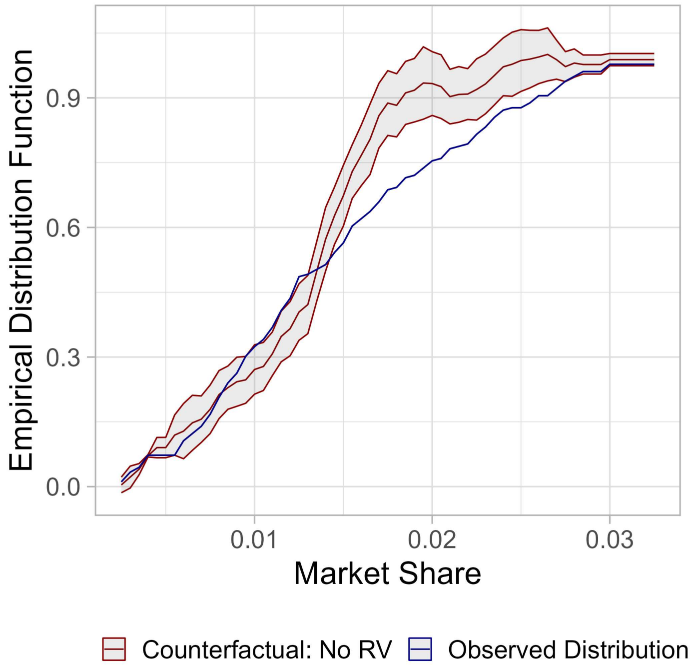


Figure 4.A.14: Distributional Effect of RV Policy - Market Shares

Note: Based on distribution regression approach by Chernozhukov et al. (2013). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.

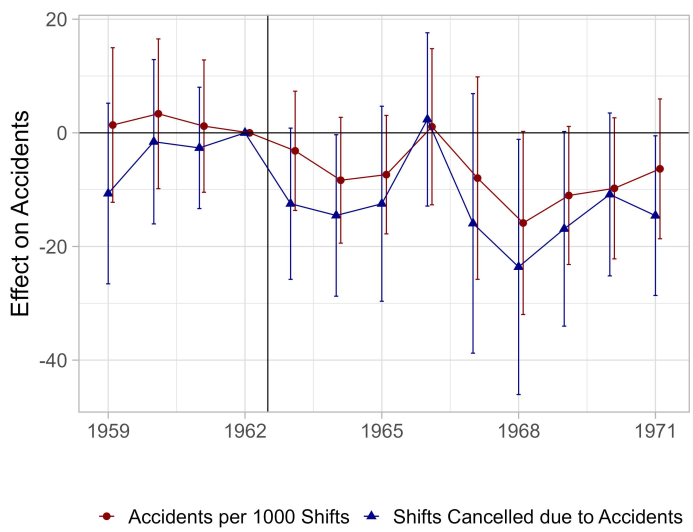


Figure 4.A.15: Effect of RV on Accidents

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.

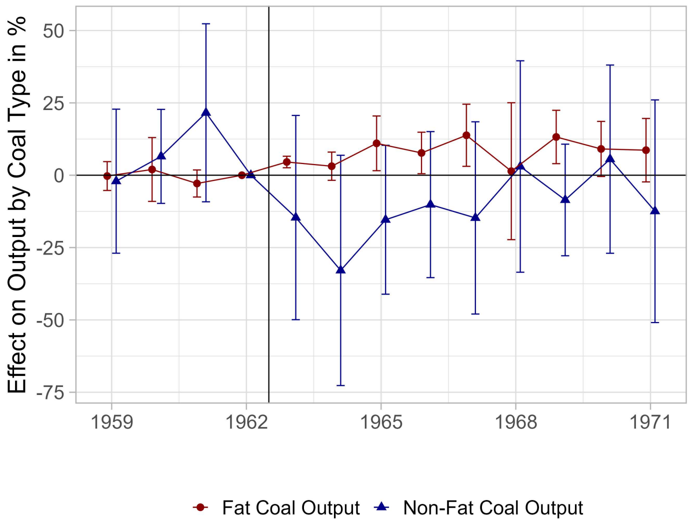
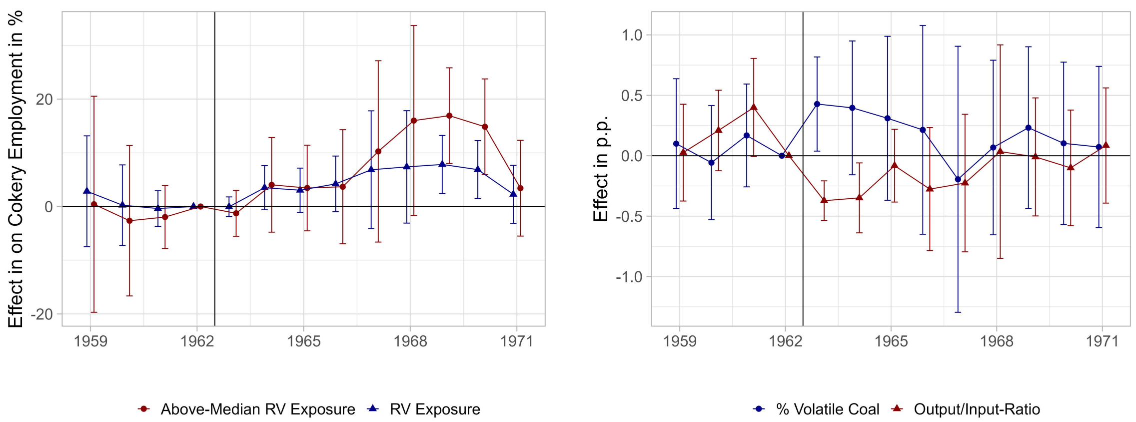


Figure 4.A.16: Effect on Share of Production by Coal Types

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence bands are reported.



(a) Cokery Employment

(b) Cokery Efficiency

Figure 4.A.17: Effect on Cokery Employment

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Only mines with cokeries over the full sample period (or until mine exit) included (intensive margin). Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.

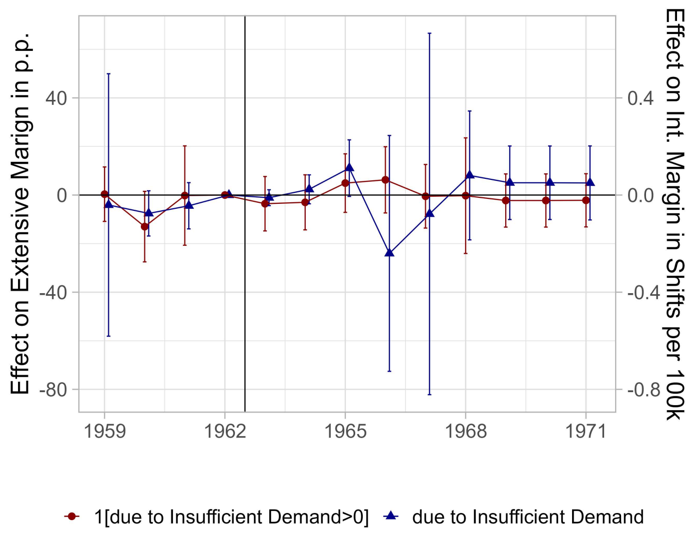


Figure 4.A.18: Effect on Cancelled Shifts due to Insufficient Demand

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.

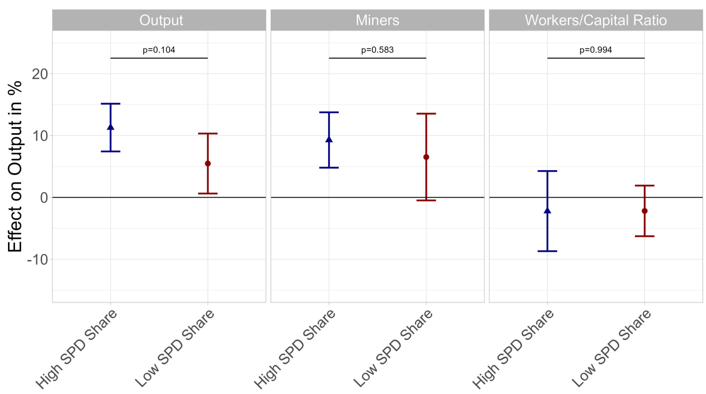


Figure 4.A.19: Heterogeneity by Political Attiude

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Sample is grouped into low and high SPD share at the median, county-level SPD share from the federal election in 1961.

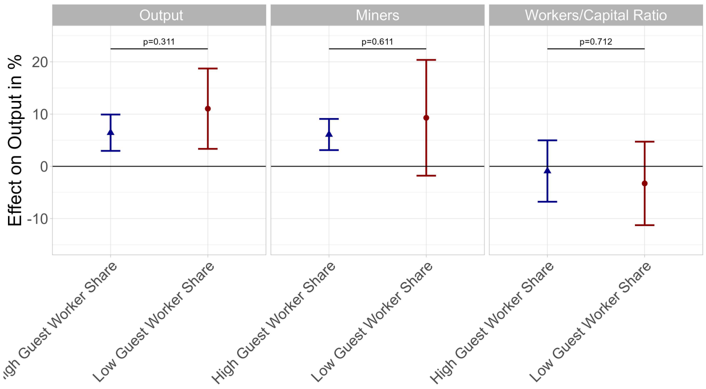


Figure 4.A.20: Heterogeneity by *Gastarbeiter* Share in Workforce

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Sample is grouped into low and high share of freign workers at the median, mine-level share of foreign workers in 1965, the first year the data is available.

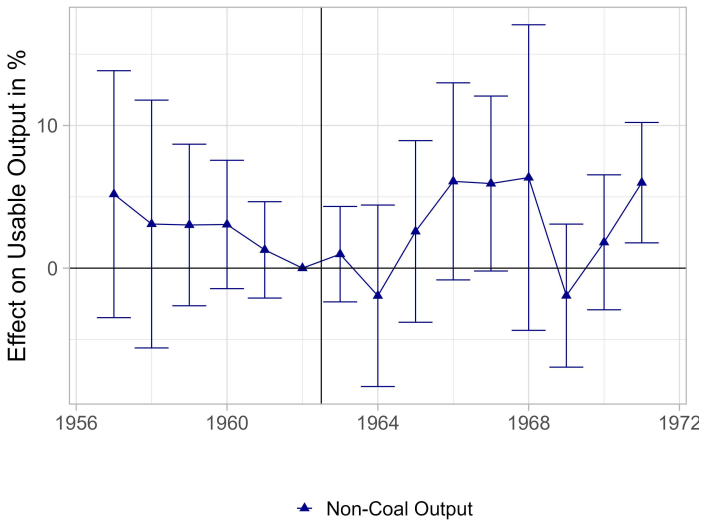
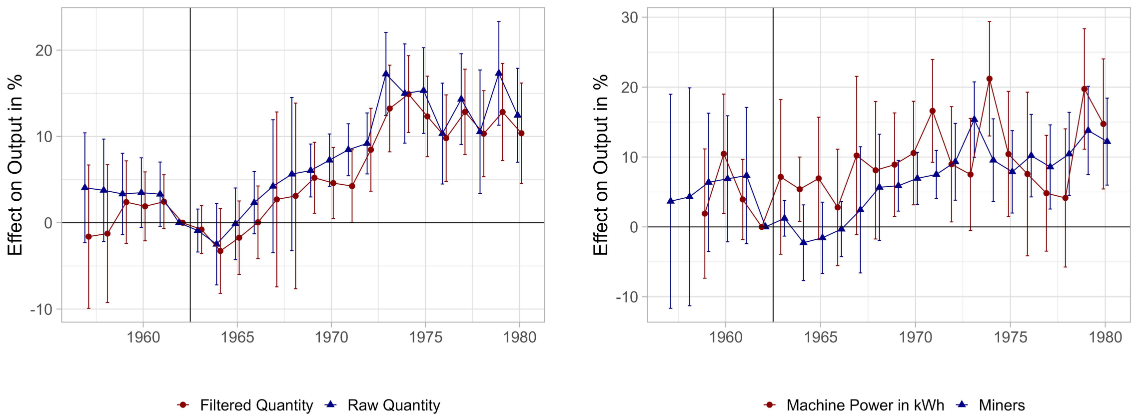


Figure 4.A.21: Effect on Non-Coal Output Quarried

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.



(a) Output

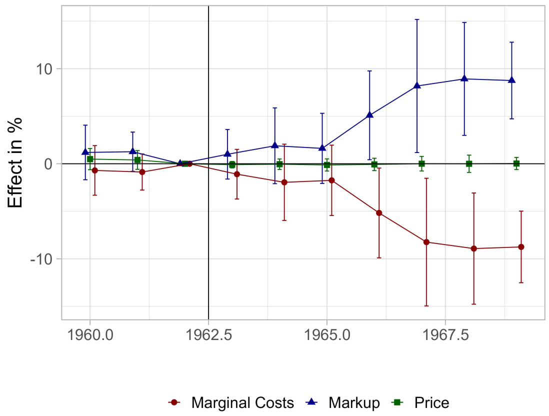
(b) Inputs

Figure 4.A.22: Sample until 1980

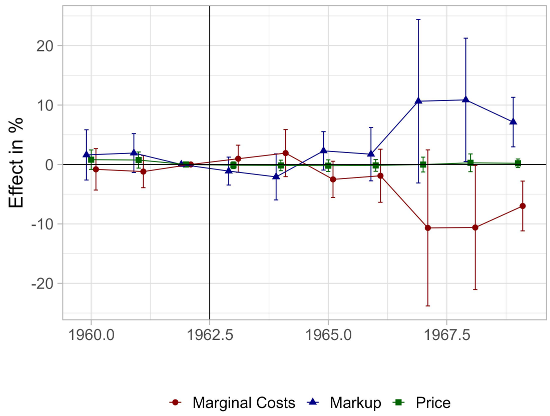
Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Only mines included which have been in operation throughout the full period 1957 to 1980. Mines are aggregated to their 1980 version in case they merged over time.

Figure 4.A.23: Effect on Wages

Note: Based on equation (4.9). Coefficients multiplied with mean [$RV\ Exposure_j$] for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.



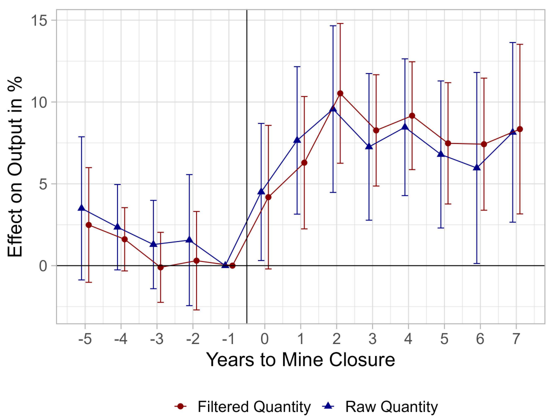
(a) All Mines



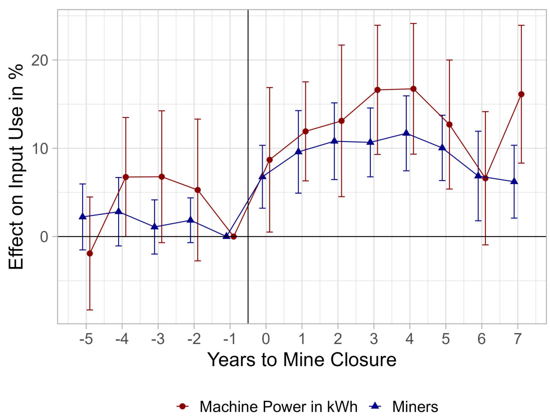
(b) Only Mines with Own Power Plants

Figure 4.A.24: Markups based on Electricity as Material with Identical Input Prices

Note: Based on equation (4.9). Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.



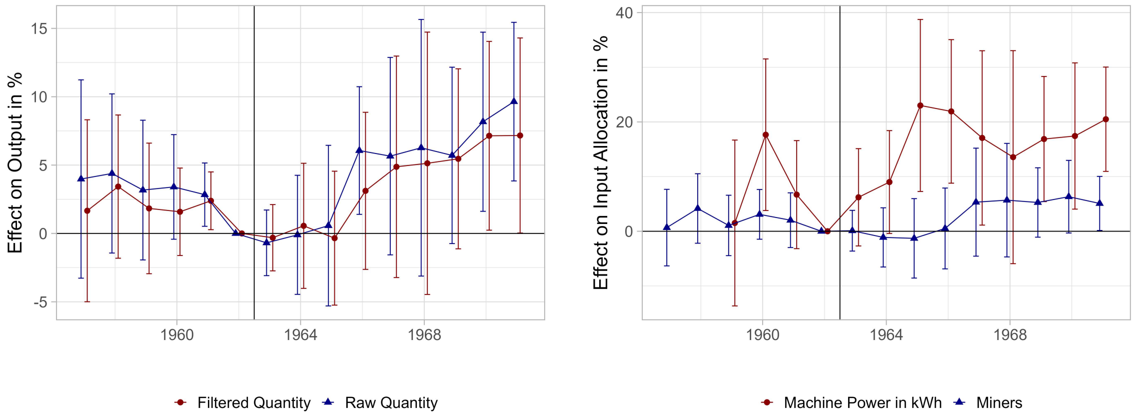
(a) Output



(b) Inputs

Figure 4.A.25: Event Study Estimates

Note: This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported.

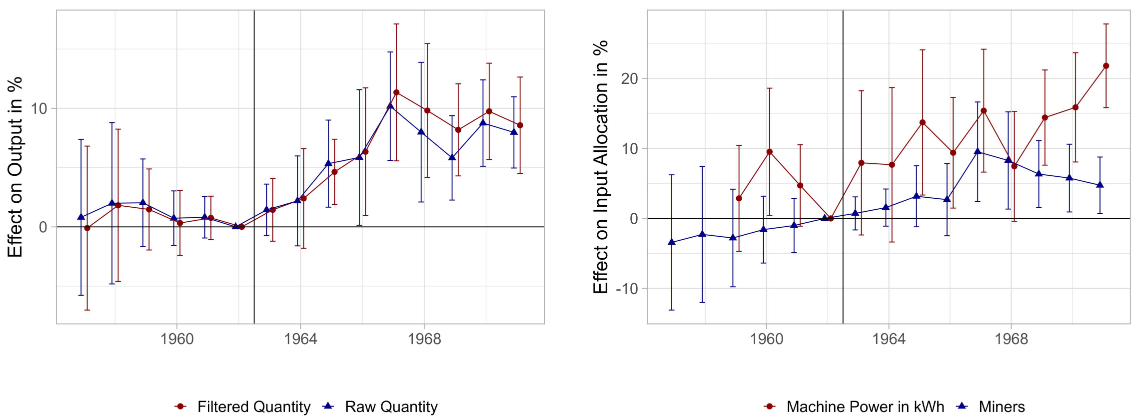


(a) Output

(b) Inputs

Figure 4.A.26: Binary Treatment (Extensive Margin)

Note: Based on equation (4.9) with binary exposure. This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Treated stations are those with positive $[RV\ Exposure_j]$.

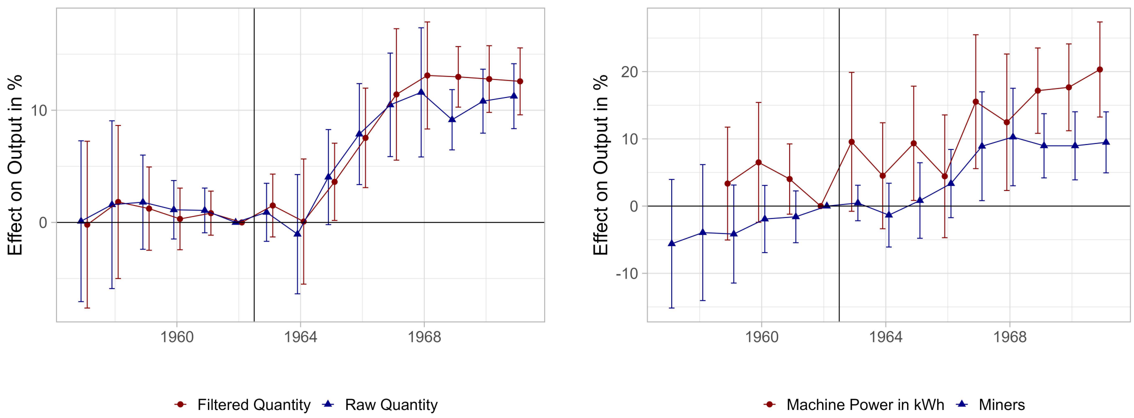


(a) Output

(b) Inputs

Figure 4.A.27: Balanced Sample

Note: Based on equation (4.9). This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Only mines included which have been in operation throughout the full period 1957 to 1971.



(a) Output

(b) Inputs

Figure 4.A.28: Sample based on only Multi-Mine Firms

Note: Based on equation (4.9). This plot documents the effect of the policy uptake on output and input usage of the remaining mines of the same firm. Coefficients multiplied with mean $[RV\ Exposure_j]$ for strictly positively treated mines to give an effect for mean exposure for treated mines. Standard errors are clustered at the firm/owner level and 90% confidence intervals are reported. Only mines included which have been owned by a firm that had at least two mines operating in the pre-policy year 1962.

Appendix B: Additional Tables

Table 4.B.1: Elasticities of Demand Over Time

	ln(Coal Output)			
	(1)	(2)	(3)	(4)
Panel A: OLS				
ln(German Coal Price Index)	-1.186^{***}			
	(0.147)			
ln(Oil Price Index)		0.483^{**}		
		(0.114)		
ln(German Coal Price Index)$\times$1[Year$\leq$1958]			-0.114^{***}	
			(0.003)	
ln(German Coal Price Index)$\times$1[Year>1958]			-1.139^{**}	
			(0.262)	
ln(Oil Price Index)$\times$1[Year$\leq$1958]				-0.072
				(0.055)
ln(Oil Price Index)$\times$1[Year>1958]				0.314
				(0.279)
Panel B: Instrumental Variable				
ln(German Coal Price Index)	-1.935^{***}			
	(0.461)			
ln(Oil Price Index)		0.685^{***}		
		(0.141)		
ln(German Coal Price Index)$\times$1[Year$\leq$1958]			-0.665^{***}	
			(0.019)	
ln(German Coal Price Index)$\times$1[Year>1958]			-2.099^{***}	
			(0.232)	
ln(Oil Price Index)$\times$1[Year$\leq$1958]				0.026
				(0.083)
ln(Oil Price Index)$\times$1[Year>1958]				0.849^{***}
				(0.224)
F-Statistic IV	21.19	180.09	21.19	180.09
Observations	17	17	17	17

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data on years 1955 to 1972. In third row, dummy for years after 1958 is not reported.

Table 4.B.2: Descriptive Statistics

Variable	Time Span	N	Mean	SD
RV Uptake				
$1[\text{Closed via } RV]_i$	1956-1971	1,512	0.215	0.411
$1[RV]_i$	1956-1971	1,512	0.380	0.485
Production Data				
Raw Extraction$_{it}$ (in 1000 tonnes)	1957-1971	1,291	1,993.9	1,082.9
Coal Production$_{it}$ (in 1000 tonnes)	1956-1971	1,512	1,252.5	763.7
Workers$_{it}$	1956-1971	1,512	3,549.8	2,178.1
Miners$_{it}$	1956-1971	1,502	2,305.1	1,405.8
Machine Power$_{it}$ (in kWh)	1959-1971	1,060	6,451.1	4,099.2
Electricity Usage$_{it}$ (in kWh)	1959-1969	960	81,821.7	51,368.4
Mine Characteristics				
Conveyor Tunnels$_{it}$	1959-1971	1,069	1.731	0.816
Depth of Mine$_{it}$ (in m)	1959-1971	1,078	913.7	208.9
Coal Layer Thickness$_{it}$ (in m)	1956-1971	1,333	125.4	28.8
% Coal Angle Up to 40 Degrees$_{it}$	1959-1971	1,085	81.22	26.09
Technology Adoption				
Mining Points$_{it}$	1957-1971	1,291	12.94	9.59
Mechanized Mining Points$_{it}$	1959-1971	1,085	5.42	3.66
% Mechanized Production$_{it}$	1959-1971	1,072	65.53	34.00
Others				
Wages$_{it}$ (in DM/Shift)	1957-1969	1,081	31.95	8.37
Wages Miners$_{it}$ (in DM/Shift)	1957-1969	1,081	34.54	9.01
% Shifts Cancelled Due to Insufficient Demand$_{it}$	1957-1969	1,081	0.77	1.60
% Shifts Cancelled Due to Reconstruction$_{it}$	1957-1969	1,081	0.44	0.33
Construction Speed$_{it}$ (cm/Day)	1957-1971	1,284	161.68	75.96
Water Inflow$_{it}$ (in m^3)	1959-1971	1,054	1,925.9	3,031.2

Note: Data is aggregated to mines as of 1971, the end of the panel, to account for mergers throughout the sample period.

Table 4.B.3: Production Function Estimation

	Baseline		Robustness	
Production Function:	Leontief		Cobb-Douglas	
Material:	Pit Wood		Electricity	
	(OLS)	(PFA)	(OLS)	(PFA)
$\hat{\beta}_L$	0.794	0.801	0.677	0.696
	(0.030)	(0.181)	(0.054)	(0.105)
$\hat{\beta}_K$	0.175	0.106	0.158	0.103
	(0.047)	(0.052)	(0.025)	(0.041)
$\hat{\beta}_M$			0.149	0.119
			(0.034)	(0.087)
Scale		0.907		0.918
		(0.190)		(0.116)
Median Markup		1.148		1.029
		(0.253)		(0.196)
Observations	922	922	798	798

Note: Standard errors are block-bootstrapped with 100 repetitions for PFA. Standard errors clustered at mine level for OLS.

Table 4.B.4: Selection into *RV* Exposure

	1[Owner is *RV* Uptaker]$_i$		[*RV Exposure$_j$*]	
Production Measures				
Standardized TFP	0.065	0.066	0.103*	0.103*
	(0.105)	(0.108)	(0.056)	(0.057)
log(Coal Production)	0.267	0.369	−0.157	−0.122
	(0.293)	(0.360)	(0.118)	(0.140)
log(Miners)	−0.035	−0.196	0.171	0.123
	(0.313)	(0.328)	(0.139)	(0.151)
log(Machine Power)	−0.064	−0.088	−0.114	−0.122
	(0.178)	(0.190)	(0.085)	(0.090)
log(Mining Points)	−0.159	−0.088	0.008	0.026
	(0.193)	(0.193)	(0.069)	(0.070)
% Mechanized Production	−0.001	−0.0003	0.001	0.001
	(0.003)	(0.003)	(0.001)	(0.001)
Mine Characteristics				
Mergel Depth	0.0006*	0.0006*	−0.0001	−0.0001
	(0.0003)	(0.0003)	(0.0001)	(0.0001)
log(Historical Coal Layer Thickness)	0.103	0.198	0.104	0.129
	(0.461)	(0.481)	(0.187)	(0.194)
% Coal Layers up to 25 Degrees	0.002	0.002	0.002	0.002
	(0.003)	(0.003)	(0.001)	(0.001)
log(Year of Mine Foundation)	−1.302	−1.831	−1.817	−1.943
	(4.917)	(4.797)	(1.637)	(1.669)
log(Coal Layer Thickness)	−0.244	−0.263	−0.020	−0.023
	(0.344)	(0.360)	(0.157)	(0.170)
Coal Type				
% Lean Coal	0.003	0.004	−0.002	−0.001
	(0.003)	(0.004)	(0.002)	(0.002)
% Fat Coal	−0.002	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.001)	(0.001)
% Anthracite Coal	0.003	0.002	−0.001	−0.001
	(0.002)	(0.003)	(0.001)	(0.001)
% Gas Coal	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.002)	(0.001)	(0.001)
Firm Characteristics				
ln(Number of Mines)	0.667***	0.659***	0.126***	0.124***
	(0.075)	(0.075)	(0.028)	(0.028)
Coal District FE	No	Yes	No	Yes
Observations	67			

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Omitted category in share of coal production is charcoal. Values as of 1962, the pre-policy year. Standard errors clustered at firm/owner level ($N = 35$). Small mines ('Kleinzechen') included. Coefficient of intercept in model (1) and (3) not reported.

Table 4.B.5: Robustness Checks: Extensive Margin

	1[Closure via RV]$_i$					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
$\log(\hat{c}_i)$	1.352***	1.365***	1.251***			
	(0.176)	(0.294)	(0.316)			
Standardized TFP$_i$				see baseline results Table 1		
1[High Cancelled Shifts]$_i$			0.456***			
			(0.118)			
1[Closure of Same Coal Type Mine]$_i$			−0.489***			
			(0.099)			
Panel B: Instrumental Variable						
$\log(\hat{c}_i)$	2.175***	2.802***	2.192***			
	(0.443)	(0.725)	(0.539)			
Standardized TFP$_i$				−0.749***	−0.247*	−0.126
				(0.173)	(0.129)	(0.159)
1[High Cancelled Shifts]$_i$			0.456***	0.195	0.370**	0.412**
			(0.108)	(0.200)	(0.159)	(0.160)
1[Closure of Same Coal Type Mine]$_i$			−0.434***	0.108	−0.342**	−0.451***
			(0.142)	(0.196)	(0.153)	(0.162)
F-Statistic First Stage	*22.89*	*18.90*	*16.41*	*38.07*	*10.70*	*8.90*
IV	Pooled IV	Pooled IV	Pooled IV	Coal Angle	Marl Thickness	Seam Thickness
Mining District FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm/Owner FE	No	Yes	Yes	Yes	Yes	Yes
Observations	96	96	96	96	96	96

Note: Significance levels of 10%, 5% and 1% are denoted by *, ** and ***. Marginal costs is averaged over 1960-1961 (no data for 1959). All individual IVs standardized. Standard errors are clustered at the firm/owner level.

Table 4.B.6: Non-Linear Treatment Effects

	ln(Output)	ln(Machinery)	ln(Miners)
Panel A: Below/Above Median			
$[Low\ RV\ Exposure_j] \times 1[Year > 1965_t]$	0.094	0.345	0.095
	(0.193)	(0.592)	(0.219)
$[High\ RV\ Exposure_j] \times 1[Year > 1965_t]$	0.595***	0.790**	0.722***
	(0.151)	(0.293)	(0.129)
Panel B: Quadratic Relationship			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	−0.126	−0.117	−0.315
	(0.260)	(0.390)	(0.260)
$[RV\ Exposure_j]^2 \times 1[Year > 1965_t]$	0.760**	0.839	1.055***
	(0.362)	(0.650)	(0.351)
Mine FE	Yes	Yes	Yes
Coal District - Year FE	Yes	Yes	Yes
Observations	1,012	861	1,012

Note: Based on equation (4.9) with adapted treatment variable and pooled post-dummy with years after the majority of exits took place, i.e., 1965.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at firm/owner level.

Table 4.B.7: Production Function: Robustness Checks

	Standardized TFP				St. LP	Standardized TFP	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Production Function	Leontief	C.-D.	Translog	Leontief	LP	Leontief	Leontief
Material Input	Pit wood	Electric.	Pit wood	Pit wood	-	Pit wood	Pit wood
Sample Period	'59-'71	'59-'69	'59-'71	'59-'80	'59-'80	'59-'71	'59-'71
Other Change						Time-Variant PF	Mechanization
Panel A: Effect after 1965							
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.677***	0.695***	0.441*	0.497**	0.486***	0.139	0.353
	(0.248)	(0.193)	(0.248)	(0.227)	(0.124)	(0.090)	(0.237)
Panel B: Effect after 1967							
$[RV\ Exposure_j] \times 1[Year > 1967_t]$	1.062***	1.037***	0.742***	0.238*	0.365***	0.236***	0.413***
	(0.129)	(0.137)	(0.169)	(0.149)	(0.129)	(0.045)	(0.129)
Mine FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Coal District - Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	922	798	922	1,064	1,289	922	922

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at firm/owner level.

Table 4.B.8: Production Function: Robustness Checks Exit

	1[Closure via RV]$_i$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Production Function	Cobb-Douglas		Translog		Leontief		Leontief	
Material Input	Electricity		Pit Wood		Pit Wood		Pit Wood	
Other Change					Time-Var. PF		Mechanization	
Panel A: OLS								
Standardized TFP$_i$	-0.138^{***}	-0.214^{***}	-0.114^{**}	-0.169^{***}	-0.054	-0.105^{**}	-0.082^{*}	-0.081^{**}
	(0.048)	(0.065)	(0.045)	(0.047)	(0.044)	(0.047)	(0.049)	(0.040)
Panel B: IV								
Standardized TFP$_i$	-0.281^{***}	-0.387^{***}	-0.349^{***}	-0.394^{***}	-0.465^{***}	-0.557^{***}	-0.333^{***}	-0.411^{***}
	(0.063)	(0.116)	(0.081)	(0.100)	(0.131)	(0.154)	(0.078)	(0.147)
F-Statistic First Stage	*41.16*	*24.37*	*33.44*	*27.20*	*15.92*	*13.89*	*29.38*	*10.30*
Owner FE	No	Yes	No	Yes	No	Yes	No	Yes
Coal District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	96	96	96	96	96	96	96	96

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at firm/owner level.

Table 4.B.9: Identification and Inference

	log(Output)	log(Miners)	log(Machinery)
	(1)	(2)	(3)
Panel A: Baseline			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.318^{***}	0.299^{**}	0.376^{***}
(cluster firm)	(0.100)	(0.111)	(0.139)
(cluster firm 1962)	(0.141)	(0.167)	(0.154)
(cluster mine)	(0.120)	(0.137)	(0.149)
(Conley spatial - 12km)	(0.106)	(0.099)	(0.131)
Panel B: Year FE instead of Coal District-Year FE			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.288^{***}	0.290^{**}	0.309^{*}
	(0.107)	(0.109)	(0.166)
Panel C: Region- instead of Coal District-Year FE			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.277^{***}	0.227^{**}	0.285^{**}
	(0.090)	(0.088)	(0.126)
Observations	1,012	1,012	861

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Significance levels of 10%, 5% and 1% are denoted by *, ** and ***.

Table 4.B.10: Analysis Weighted by Mine Size

	log(Output) (1)	log(Miners) (2)	log(Machinery) (3)
Panel A: Baseline			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.269***	0.233**	0.324**
	(0.093)	(0.101)	(0.135)
Panel B: Weighted by Output in 1962 (Mine Size)			
$[RV\ Exposure_j] \times 1[Year > 1965_t]$	0.365***	0.309***	0.373***
	(0.075)	(0.075)	(0.118)
Mine FE	Yes	Yes	Yes
Coal District-Year FE	Yes	Yes	Yes
Observations	970	970	834

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Only observations of mines which operated in the pre-policy year 1962 included (to have a weight). Hence, baseline results slightly change. Significance levels of 10%, 5% and 1% are denoted by *, ** and ***.

Table 4.B.11: Merger Effects

	log(Output)	log(Miners)	log(Machinery)	Stand. LP	Standardized TFP Pit Wood	Standardized TFP Electricity	1[Survival]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1[Post-Merger]	−0.065	−0.085**	0.027	0.034	−0.036	−0.069	−0.016
	(0.040)	(0.032)	(0.055)	(0.061)	(0.121)	(0.096)	(0.085)
Observations	1,012	1,012	861	1,012	778	660	4,544

Note: Based on equation (4.9) with pooled post-dummy with years after the majority of exits took place, i.e., 1965. Significance levels of 10%, 5% and 1% are denoted by *, ** and ***. I control for the *RV* policy by adding a $[RV\ Exposure_j] \times 1[Year > 1965_t]$ interaction to the regression but omit the output.

Bibliography

Ackerberg, D. A., Caves, K. and Frazer, G. (2015), 'Identification Properties of Recent Production Function Estimators', *Econometrica* **83**(6), 2411–2451.

Adamopoulos, T., Brandt, L., Leight, J. and Restuccia, D. (2022), 'Misallocation, Selection, and Productivity: A Quantitative Analysis with Panel Data from China', *Econometrica* **90**(3), 1261–1282.

Aghion, P., Cai, J., Dewatripont, M., Du, L., Harrison, A. and Legros, P. (2015), 'Industrial Policy and Competition', *American Economic Journal: Macroeconomics* **7**(4), 1–32.

Allcott, H., Collard-Wexler, A. and O'Connell, S. D. (2016), 'How Do Electricity Shortages Affect Industry? Evidence from India', *American Economic Review* **106**(3), 587–624.

Allied Higher Commission (2019), 'Gesetz N°27 hinsichtlich der Umgestaltung des deutschen Kohlenbergbaues und der deutschen Stahl- und Eisenindustrie'.

Amiti, M., Itskhoki, O. and Knonigs, J. (2019), 'International Shocks, Variable Markups, and Domestic Prices', *Review of Economic Studies* **86**(6), 2356–2402.

Amiti, M. and Khandelwal, A. K. (2013), 'Import Competition and Quality Upgrading', *Review of Economics and Statistics* **95**(2), 476–490.

Arnold, J. M. and Javorcik, B. S. (2009), 'Gifted Kids or Pushy Parents? Foreign Direct Investment and Plant Productivity in Indonesia', *Journal of International Economics* **79**(1), 42–53.

Asker, J., Collard-Wexler, A. and De Loecker, J. (2019), '(Mis)Allocation, Market Power, and Global Oil Extraction', *American Economic Review* **109**(4), 1568–1615.

Asker, J., Collard-Wexler, A. and De Loecker, J. (2020), 'The Welfare Impact of Market Power: The OPEC Cartel'. Working Paper.

Autor, D., Dorn, D., Hanson, G. and Majlesi, K. (2019), 'Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure', *American Economic Review* **110**(10), 3139–3183.

Autor, D., Dorn, D., Katz, L., Patterson, C. and Van Reenen, J. (2020), 'The Fall of the Labor Share and the Rise of Superstar Firms', *The Quarterly Journal of Economics* **135**(2), 645–709.

Avignon, R. and Guigue, E. (2023), 'Markups and Markdowns in the French Dairy Market'. Working Paper.

Backus, M. (2020), 'Why is Productivity Correlated with Competition?', *Econometrica* **88**(6), 2415–2444.

Baily, M. N., Hulten, C., Campbell, D., Bresnahan, T. and Caves, R. E. (1992), 'Productivity Dynamics in Manufacturing Plants', *Brookings Papers on Economic Activity: Microeconomics* **1992**, 187–267.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013), 'Cross-Country Differences in Productivity: The Role of Allocation and Selection', *American Economic Review* **103**(1), 305–334.

Barwick, P. J., Kalouptsidi, M. and Zahur, N. B. (2024*a*), *Industrial Policy Implementation: Empirical Evidence from China's Shipbuilding Industry*. forthcoming.

Barwick, P. J., Kalouptsidi, M. and Zahur, N. B. (2024*b*), 'Industrial Policy: Lessons from Shipbuilding', *Journal of Economic Perspectives* **38**(4), 55–80.

Becker, S. O., Egger, P. H. and Von Ehrlich, M. (2010), 'Going NUTS: The effect of EU Structural Funds on regional performance', *Journal of Public Economics* **94**(9-10), 578–590.

Behrens, K., Mion, G., Murata, Y. and Suedekum, J. (2020), 'Quantifying the Gap between Equilibrium and Optimum under Monopolistic Competition', *The Quarterly Journal of Economics* **135**(4), 2299–2360.

Bekaert, F., Hagenbruch, T., Kastl, E., Mareels, S., Van Hoey, M., Vercammen, S. and Zeumer, B. (2021), 'The Future of the European Steel Industry', *McKinsey & Company Metals & Mining Practice* .

Berbee, P., Braun, S. T. and Franke, R. (2024), 'Reversing Fortunes of German Regions, 1926–2019: Boon and Bane of Early Industrialization?', *Journal of Economic Growth* . forthcoming.

Borusyak, K., Jaravel, X. and Spiess, J. (2024), 'Revisiting Event-study Designs: Robust and Efficient Estimation', *Review of Economic Studies* **91**(6), 3253–3285.

Braguinsky, S., Ohyama, A., Okazaki, T. and Syverson, C. (2015), 'Acqusitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry', *American Economic Review* **105**(7), 2086–2119.

Brandt, L., Van Biesebroeck, J., Wang, L. and Zhang, Y. (2017), 'WTO Accession and Performance of Chinese Manufacturing Firms', *American Economic Review* **107**(9), 2784–2820.

Bundestag, D. (1959*a*), 'Gesetz über das Zollkontingent fester Brennstoffe', *Bundesgesetzblatt* **50**, 1380–1383.

Bundestag, D. (1959*b*), 'Plenarprotokoll - Deutscher Bundestag 59. Sitzung', *Plenarprotokolle Deutscher Bundestag* .

Bundestag, D. (1960), 'Gesetz zur Minderung des Minerölsteuergesetzes', *Bundesgesetzblatt* **20**, 241–242.

Bundestag, D. (1962), 'Plenarprotokoll - Deutscher Bundestag 30. Sitzung', *Plenarprotokolle Deutscher Bundestag* .

Bundestag, D. (1963), 'Gesetz zur Förderung der Rationalisierung im Steinkohlenbergbau', *Bundesgesetzblatt* **44**, 549–561.

Bundestag, D. (1965), 'Plenarprotokoll - Deutscher Bundestag 191. Sitzung', *Plenarprotokolle Deutscher Bundestag* .

Burckhardt, H. (1968), *Deutscher Steinkohlenbergbau im Spannungsfeld zwischen Politik und Wirtschaft - Eine Dokumentation*, Deutscher Industrieverlag, Köln.

Callaway, B., Goodman-Bacon, A. and Sant'Anna, P. H. (2024*a*), Difference-in-differences with a Continuous Treatment, Technical report, National Bureau of Economic Research.

Callaway, B., Goodman-Bacon, A. and Sant'Anna, P. H. (2024*b*), Event studies with a continuous treatment, *in* 'AEA Papers and Proceedings', Vol. 114, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 601–605.

Callaway, B. and Sant'Anna, P. H. (2021), 'Difference-in-differences with Multiple Time Periods', *Journal of Econometrics* **225**(2), 200–230.

Carret, V. (2023), 'Distribution of Power and Ordered Competition in the European Coal and Steel Community'. Working Paper.

Chen, G., Herrera, A. M. and Lugauer, S. (2022), 'Policy and Misallocation: Evidence from Chinese Firm-level Data', *European Economic Review* **149**, 104260.

Chernozhukov, V., Fernández-Val, I. and Melly, B. (2013), 'Inference on Counterfactual Distributions', *Econometrica* **81**(6), 2205–2268.

Cingano, F., Palomba, F., Pinotti, P. and Rettore, E. (2022), 'Making Subsidies Work: Rules vs. Discretion', *Bank of Italy Temi di Discussione (Working Paper) No* **1364**.

Collard-Wexler, A. (2013), 'Demand Fluctuations in the Ready-Mix Concrete Industry', *Econometrica* **81**(3), 1003–1037.

Commission of European Union (1988), 'Commission Regulation (EEC) No 1272/88 of 29 April 1988 Laying Down Detailed Rules for Applying the Set-aside Incentive Scheme for Arable Land'.

Commission of European Union (2016), 'Commission Delegated Regulation (EU) 2016/1612 of 8 September 2016 Providing Aid for Milk Production Reduction'.

Conley, T. G. (1999), 'GMM Estimation with Cross sectional Dependence', *Journal of Econometrics* **92**(1), 1–45.

Council of European Union (2006), 'Council Regulation (EC) No 1198/2006 of 27 July 2006 on the European Fisheries Fund'.

Criscuolo, C., Martin, R., Overman, H. G. and Van Reenen, J. (2019), 'Some Causal Effects of an Industrial Policy', *American Economic Review* **109**(1), 48–85.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020), 'Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects', *American Economic Review* **110**(9), 2964–2996.

De Loecker, J. (2011), 'Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity', *Econometrica* **79**(5), 1407–1451.

De Loecker, J. (2013), 'Detecting Learning by Exporting', *American Economic Journal: Microeconomics* **5**(3), 1–21.

De Loecker, J. and Eeckhout, J. (2018), 'Global Market Power', *NBER Working Paper 24768* .

De Loecker, J., Eeckhout, J. and Unger, G. (2020), 'The Rise of Market Power and the Macroeconomic Implications', *Quarterly Journal of Economics* **135**(2), 561–644.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K. and Pavcnik, N. (2016), 'Prices, Markups, and Trade Reform', *Econometrica* **84**(2), 445–510.

De Loecker, J. and Scott, P. T. (2022), 'Markup Estimation Using Production and Demand Data - An Application to the US Brewing Industry'. Working Paper.

De Loecker, J. and Warzynski, F. (2012), 'Markups and Firm-Level Export Status', *American Economic Review* **102**(6), 2437–2471.

Dechezleprêtre, A., Díaz, L., Fadic, M. and Lalanne, G. (2023), 'How the Green and Digital Transitions are Reshaping the Automotive Ecosystem'.

Delabastita, V. and Rubens, M. (2023), 'Colluding Against Workers'. Working Paper.

Der Spiegel (1958*a*), 'Die Schnüffel-Kommission', *Der Spiegel* **2**, 23–24.

Der Spiegel (1958*b*), 'Die Öldruck-Bremse', *Der Spiegel* **2**, 23–24.

Der Spiegel (1963), 'Zwei statt drei', *Der Spiegel* **3**.

Dippel, C., Gold, R., Heblich, S. and Pinto, R. (2022), 'The Effect of Trade on Workers and Voters', *The Economic Journal* **132**(641), 199–217.

Ehrlich, M. v. and Seidel, T. (2018), 'The Persistent Effects of Place-based Policy: Evidence from the West-German Zonenrandgebiet', *American Economic Journal: Economic Policy* **10**(4), 344–374.

Federal Statistical Office of Germany (1960), 'Statistisches Jahrbuch für die Bundesrepublik Deutschland 1960', *Statistisches Jahrbuch für die Bundesrepublik Deutschland* .

Federal Statistical Office of Germany (1965), 'Statistisches Jahrbuch für die Bundesrepublik Deutschland 1964', *Statistisches Jahrbuch für die Bundesrepublik Deutschland* .

Fritzsche, M. and Wolf, N. (2023), 'Fickle Fossils. Economic Growth, Coal and the European Oil Invasion, 1900-2015', *CESifo Working Papers 10805* .

Fudenberg, D. and Tirole, J. (1986), 'A Theory of Exit in Duopoly', *Econometrica* **54**(4), 943–960.

Gagliardi, L., Moretti, E. and Serafinelli, M. (2023), The World's Rust Belts: The Heterogeneous Effects of Deindustrialization on 1,993 Cities in Six Countries, Technical report, National Bureau of Economic Research.

Gandhi, A., Navarro, S. and Rivers, D. A. (2020), 'On the Identification of Gross Output Production Functions', *Journal of Political Economy* **128**(8), 2973–3016.

Garin, A. and Rothbaum, J. L. (2024), The Long-Run Impacts of Public Industrial Investment on Local Development and Economic Mobility: Evidence from World War II, Technical report, National Bureau of Economic Research.

Gatzka, W. (1996), *"Der Hüttenvertrag - Die Ablösung eines über hundert Jahre gewachsenen Verbundes durch einen Vertrag"*, Shaker Verlag, Aachen.

Gebhardt, G. (1957), *Ruhrbergbau: Geschichte, Aufbau und Verflechtung seiner Gesellschaften und Organisationen*, Verlag Glückauf.

Geitling (1956), 'Gesellschaftsvertrag der "Geitling" Ruhrkohlen-Verkaufsgesellschaft mit beschränkter Haftung'.

Geitling (1960), 'Gesellschaftsvertrag der "Geitling" Ruhrkohlen-Verkaufsgesellschaft mit beschränkter Haftung'.

German Federal Commissioner for Coal (1968), 'Untersuchung über die Entwicklung der Unternehmensgrößén im deutschen Steinkohlenbergbau'.

German Federal Ministry for Economic Affairs and Climate Action (2019), ‘Abschlussbericht der Kommission Wachstum, Strukturwandel und Beschäftigung’.

Ghemawat, P. and Nalebuff, B. (1985), ‘Exit’, *RAND Journal of Economics* **16**(2), 184–194.

Ghemawat, P. and Nalebuff, B. (1990), ‘The Devolution of Declining Industries’, *Quarterly Journal of Economics* **105**(1), 167–186.

Gibson, J. K. and Harris, R. I. D. (1996), ‘Trade Liberalisation and Plant Exit in New Zealand Manufacturing’, *Review of Economics and Statistics* **78**(3), 521–529.

Grieco, P., Pinske, J. and Slade, M. (2018), ‘Brewed in North America: Mergers, Marginal Costs, and Efficiency’, *International Journal of Industrial Organization* **59**, 24–65.

Griliches, Z. and Regev, H. (1995), ‘Firm Productivity in Israeli Industry 1979–1988’, *Journal of Econometrics* **65**(1), 175–203.

Guanziroli, T. (2022), ‘Does Labor Market Concentration Decrease Wages? Evidence from a Retail Pharmacy Merger’.

Hahn, N. (2024), Who is in the Driver’s Seat? Markups, Markdowns, and Profit sharing in the Car Industry, Technical report, ZEW Discussion Papers.

Harris, R., Keay, I. and Lewis, F. (2015), ‘Protecting Infant Industries: Canadian Manufacturing and the National policy, 1870–1913’, *Explorations in Economic History* **56**, 15–31.

Heblich, S., Seror, M., Xu, H. and Zylberberg, Y. (2022), Industrial Clusters in the Long run: Evidence from Million-Rouble Plants in China, Technical report, National Bureau of Economic Research.

Heim, S., Hüschelrath, K., Schmidt-Dengler, P. and Strazzeri, M. (2017), ‘The Impact of State Aid on the Survival and Financial Viability of Aided Firms’, *European Economic Review* **100**, 193–214.

High Authority (1956*a*), ‘Decision No. 5/56’, *Amtsblatt - Hohe Behörde* **31**, 29–42.

High Authority (1956*b*), ‘Decision no. 8/56’, *Amtsblatt - Hohe Behörde* **6**, 70–80.

High Authority (1956*c*), ‘Hohe Behörde’, *Monatliches Mitteilungsblatt* **9**, 29–31.

High Authority (1965), ‘Decision no. 17/65’, *Amtsblatt der Europäischen Gemeinschaften* **65**, 3249–3254.

Hsieh, C.-T., Hurst, E., Jones, C. I. and Klenow, P. J. (2019), ‘The Allocation of Talent and U.S. Economic Growth’, *Econometrica* **87**(5), 1439–1474.

Hsieh, C.-T. and Klenow, P. J. (2009), ‘Misallocation and Manufacturing TFP in China and India’, *Quarterly Journal of Economics* **129**(4), 1403–1438.

Hsieh, C.-T. and Klenow, P. J. (2014), ‘The Life Cycle of Plants in India and Mexico’, *Quarterly Journal of Economics* **129**(3), 1035–1084.

Hsieh, C.-T. and Klenow, P. J. (2018), ‘The Reallocation Myth’. Working Paper.

Hünermund, P., Schmidt-Dengler, P. and Takahashi, Y. (2015), ‘Entry and Shakeout in Dynamic Oligopoly’. Working Paper.

Juhász, R. (2018), ‘Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade’, *American Economic Review* **108**(11), 3339–3376.

Juhász, R., Lane, N., Oehlsen, E. and Pérez, V. C. (2022), 'The Who, What, When, and How of Industrial Policy: A Text-based Approach', *What, When, and How of Industrial Policy: A Text-Based Approach (August 15, 2022)* .

Juhász, R., Lane, N. and Rodrik, D. (2023), 'The New Economics of Industrial Policy', *Annual Review of Economics* **16**.

Juhász, R. and Steinwender, C. (2023), 'Industrial policy and the Great Divergence', *Annual Review of Economics* **16**.

Khandelwal, A. K., Schott, P. K. and Wei, S.-J. (2013), 'Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporters', *American Economic Review* **103**(6), 2169–2195.

Klepper, S. and Simons, K. L. (2000), 'The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the US Tire Industry', *Journal of Political Economy* **108**(4), 728–760.

Klepper, S. and Simons, K. L. (2005), 'Industry Shakeout and Technological Change', *International Journal of Industrial Organization* **23**, 23–43.

Klepper, S. and Thompson, P. (2006), 'Submarkets and the Evolution of Market Structure', *RAND Journal of Economics* **37**(4), 861–886.

Kline, P. and Moretti, E. (2014), 'Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority', *The Quarterly journal of Economics* **129**(1), 275–331.

Lane, N. (2022), 'Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea'. Working Paper.

LaPoint, C. and Sakabe, S. (2021), 'Place-based Policies and the Geography of Corporate Investment'. Working Paper.

Levinsohn, J. and Petrin, A. (2003), 'Estimating Production Functions Using Inputs to Control for Unobservables', *Review of Economic Studies* **70**(2), 317–341.

Lieberman, M. B. (1990), 'Exit from Declining Industries: 'Shakeout' or 'Stakeout'', *RAND Journal of Economics* **21**(4), 538–554.

Lileeva, A. and Trefler, D. (2010), 'Improved Access to Foreign Markets Raises Plant-Level Productivity... For Some Plants', *Quarterly Journal of Economics* **125**(3), 1051–1099.

Lu, Y. and Yu, L. (2015), 'Trade Liberalization and Markup Dispersion: Evidence from China's WTO Accession', *American Economic Journal: Applied Economics* **7**(4), 221–253.

Manelici, I. and Pantea, S. (2021), 'Industrial Policy at Work: Evidence from Romania's Income Tax Break for Workers in IT', *European Economic Review* **133**, 103674.

Melitz, M. J. and Polanec, S. (2015), 'Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit', *RAND journal of Economics* **46**(2), 362–375.

Midrigan, V. and Xu, D. Y. (2014), 'Finance and Misallocation: Evidence from Plant-Level Data', *American Economic Review* **104**(2), 422–458.

Miller, N. H. and Weinberg, M. C. (2017), 'Understanding the Price Effects of the MillerCoors Joint Venture', *Econometrica* **85**(6), 1763–1791.

Morlacco, M. (2019), 'Market Power in Input Markets: Theory and Evidence from French Manufacturing'. Working Paper.

Olley, G. S. and Pakes, A. (1996), 'The Dynamics of Productivity in the Telecommunications Equipment Industry', *Econometrica* **64**(6), 1263–1297.

Orr, S. and Tabari, M. (2024), 'Decomposing the Within-Firm Productivity Gains from Trade: Evidence from India'. Working Paper.

Pavcnik, N. (2002), 'Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants', *Review of Economic Studies* **69**, 245–276.

Prager, E. and Schmitt, M. (2021), 'Employer Consolidation and Wages: Evidence from Hospitals', *American Economic Review* **111**(2), 397–427.

Raggi, P., Richard, G. and Ollitrault, S. (2015), Le Plan Acier de 1984 et la Crise de la Sidérurgie en Lorraine, *in* '1984-1988: les années d'alternance vues des régions', Presses Universitaires de Rennes, pp. 251–262.

Restuccia, D. and Rogerson, R. (2017), 'The Causes and Costs of Misallocation', *Journal of Economic Perspectives* **31**(3), 151–174.

Rodrik, D. (2004), 'Industrial policy for the Twenty-first Century'. Working Paper.

Rodrik, D. (2009), 'Industrial Policy: Don't ask why, Ask how', *Middle East development journal* **1**(1), 1–29.

Rubens, M. (2023*a*), 'Management, Productivity, and Technology Choices: Evidence from U.S. Mining Schools', *RAND Journal of Economics* **54**(1), 165–186.

Rubens, M. (2023*b*), 'Market Structure, Oligopsony Power, and Productivity', *American Economic Review* **113**(9), 2382–2410.

Rubens, M. (2023*c*), 'Oligopsony Power and Factor-Biased Technology Adoption'. Working Paper.

Schmitt, M. (2017), 'Do Hospital Mergers Reduce Costs?', *Journal of Health Economics* **52**, 74–94.

Schroeder, R. (1953), 'Westdeutsche Holzwirtschaft ohne Preisbindung', *Wirtschaftsdienst* **33**(5), 317–320.

Seidel, H.-C. (2014), 'Die Bergbaugewerkschaft und die Gastarbeiter Ausländerpolitik im Ruhrbergbau vom Ende der 1950er bis in die 1980er Jahre', *Vierteljahrshefte für Zeitgeschichte* **62**(1), 35–68.

Siegloch, S., Wehrhöfer, N. and Etzel, T. (2024), 'Spillover, Efficiency and Equity Effects of Regional Firm Subsidies', *American Economic Journal: Economic Policy* **17**(1), 144–180.

Statistisches Landesamt Nordrhein-Westfalen (1964), 'Die Industrie Nordrhein-Westfalens Jahreszahlen 1962', *Beiträge zur Statistik des Landes Nordrhein-Westfalens* **173**.

Stiebale, J., Suedekum, J. and Woessner, N. (2024), 'Robots and the Rise of European Superstar Firms', *International Journal of Industrial Organization* **94**, 103085.

Stiebale, J. and Szücs, F. (2022), 'Mergers and Market Power: Evidence from Rivals' Responses in European Markets', *RAND Journal of Economics* **53**(4), 678–702.

Stilz, D. (1969), 'Die Begünstigung des Steinkohlenbergbaus des Ruhrgebiets durch die öffentliche Finanzwirtschaft', *Journal of Contextual Economics–Schmollers Jahrbuch* (2), 151–184.

Storchmann, K. (2005), 'The Rise and Fall of German Hard Coal Subsidies', *Energy Policy* **33**(11), 1469–1492.

Sun, L. and Abraham, S. (2021), 'Estimating Dynamic Treatment Effects in Event studies with Heterogeneous Treatment Effects', *Journal of Econometrics* **225**(2), 175–199.

Syverson, C. (2004), 'Market Structure and Productivity: A Concrete Example', *Journal of Political Economy* **112**(6), 1181–1222.

Syverson, C. (2011), 'What Determines Productivity?', *Journal of Economic Literature* **49**(2), 326–365.

Takahashi, Y. (2015), 'Estimating a War of Attrition: The Case of the US Movie Theater Industry', *American Economic Review* **105**(7), 2204–2241.

Topalova, P. and Khandelwal, A. (2011), 'Trade Liberalization and Firm Productivity: Evidence from India', *Review of Economics and Statistics* **93**(3), 995–1009.

Unternehmensverband Ruhrbergbau (1961), *Jahresbericht des Unternehmensverbandes Ruhrbergbau für die Jahre 1958 bis 1960*, Essen.

Wasserstraßen- und Schifffahrtsverwaltung des Bundes (2021), 'Verknet-bwastr', *Karten Wasserstraßen- und Schifffahrtsverwaltung des Bundes* .

Whinston, M. (1988), 'Exit with Multiplant Firms', *RAND Journal of Economics* **19**(4), 568–588.

Chapter 5

Immigration, Workforce Composition, and Organizational Performance: The Effect of Brexit on NHS Hospital Quality

Coauthor(s): Henrique Castro-Pires (Miami Herbert Business School), Marco Mello (University of Aberdeen), Giuseppe Moscelli (University of Surrey)

Abstract: Restrictive immigration policies may force firms to abruptly change their workforce composition. But how does this impact the performance of these organizations? We study the effects of the 2016 Brexit referendum, which led to a drop in the share of EU nationality nurses in English hospitals. Using high-quality administrative patient-level data and a continuous difference-in-differences design which exploits the different pre-referendum hospital exposure to the shock, we estimate the causal effect of the workforce composition changes on hospital quality of care. We find that, in the post-referendum period, emergency patients admitted to NHS hospitals with a mean pre-referendum share of EU nurses faced an increase in mortality risk, equivalent to about 1,485 additional deaths per year. These findings are consistent with a theory model that predicts a decrease in the quality of newly hired hospital workers to avert labour shortages. We provide empirical evidence in support of this mechanism by showing that the foreign joiner nurses hired in the post-referendum period were assigned to lower salary grades than those hired prior to the referendum, indicating lower levels of skills and job experience.

5.1 Introduction

The shortage of skilled labour is one of the main bottlenecks which is expected to harm the growth of Western economies in the upcoming decades. For the past few years, the US and European Union (EU) member states have been experiencing an insufficient labour supply of skilled workers in many professions such as construction, healthcare, or manufacturing, and face hiking job opening rates (U.S. Bureau of Labor Statistics, 2023, European Labour Authority, 2023, Eurostat, 2023*a*). In many countries, this pattern will persist, and likely be reinforced, in light of the currently ageing demographic trends. In the EU, the working-age population is predicted to decline by 6% until 2040 (Eurostat, 2023*b*), and OECD countries are foreseen to lose on average 10% of their working-age population by 2060 (OECD, 2021). Policymakers have been trying to tackle this critical economic challenge by acting on several policy levers, for example by increasing and facilitating labour market participation, incentivizing longer working hours or later retirement, and by improving the quality of job matches between employers and employees. Although the adoption of the aforementioned policies is usually beneficial, reducing the labour shortages in many sectors of Western economies might be still difficult without attracting migrant workers from other countries. However, immigration policies trigger an open dilemma for policymakers, who face the trade-off between increasing the labour supply needed to raise aggregate productivity at national level and the implied cost of immigration, such as the rise of populistic movements and parties (Rodrik, 2021, Guriev and Papaioannou, 2022). While there is a vast economics literature on the effects of labour supply expansions achieved through increases in the foreign labour force (e.g., Friedberg and Hunt 1995, Friedberg 2001, Card 2001, Borjas 2003, Dustmann et al. 2013, Peri 2012, 2014, Peri et al. 2015, among many others), robust evidence on the effects that immigration barriers can produce on outputs of economic interest, such as organizational performance, is much scanter.[1]

With this paper we contribute to fill this evidence gap, by investigating the effects on organizational performance and consumer outcomes caused by a change in the labour supply of skilled migrant workers. To do so, we exploit the outcome of the Brexit referendum as a persistent and large-scale, negative labour supply shock to an exceptionally tight labour market: the labour market for nurses in English public hospitals. In this market, vacancy rates have been persistently at 10% and employment of foreign-trained nurses has been 2.5 times as high as the OECD average (OECD, 2017), mainly due to a lack of domestic workers.[2] In this paper, we show that the change in the workforce composition of skilled migrant nurses joining English National Health Service (NHS) hospitals had a negative impact on the health outcomes of hospital patients.

To identify the causal effect of this labour supply shock on hospital organization performance and patient health outcomes, we exploit quasi-experimental variation in the degree of pre-

[1] As an exception, see (Lee et al., 2022).

[2] The shortage of skilled nurses and healthcare workers is not unique to England. In the EU, the occupation of 'nursing professionals' is the number one field for which most EU member states report labour shortages (European Commission, 2020).

shock exposure of NHS hospital organizations to the employment of EU nurses within their workforce and, hence, to the reduced immigration of EU nurses following the victory of the vote to leave the EU. Across English NHS hospitals, the pre-referendum share of EU nurses, among all nurses employed, ranged from 0.5 to 22%. As the Brexit vote was unanticipated, hospitals could not strategically plan for their workforce needs in advance, allowing us to compare patient health outcomes – in particular, in-hospital mortality and unplanned emergency readmissions – and other hospital organization performance indicators before and after the Brexit referendum, using a continuous treatment difference-in-differences design based on the differential exposure to the Brexit shock due to heterogenous pre-referendum workforce composition.

The English NHS institutional setting is ideal to answer this research question. Hospitals within the English NHS are subject to identical institutional regulations (such as pay agreements and financial rules) and clinical guidelines, both set at national level, making these organizations rather homogeneous in their service delivery model and, hence, more suitable to compare than firms in other productive sectors. Nevertheless, NHS hospitals have complete autonomy in the hiring decisions of clinical staff such as nurses and doctors; thus, different hospital organizations might have reacted heterogeneously to labour supply shortages of nurses and doctors. Moreover, hospitals offer highly-valuable labour-intensive services to consumers, because health care is a basic need. Finally, 'quality' of healthcare service is more objectively and unambiguously defined and measured than the quality of consumable products and durable assets, where consumer preferences' heterogeneity is greater.

The main analyses of this study are based on the linkage of three high-quality and rich datasets, covering the period from 2012 to 2019: the administrative payroll records of all English NHS hospitals provide us with information on the monthly composition of its clinical workforce; we gather patient-level mortality and readmissions indicators, as well as other covariates such as patient age, sex and comorbidities, from the universe of patient admission records at English NHS acute care hospitals; and we capture the nurses' job satisfaction with the quality of services provided to patients, by employing large-scale NHS staff surveys collected every year at the nurse level. We focus our analysis on the health outcomes of emergency hospital patients, as in the NHS they typically cannot choose which hospital to be admitted to (Gaynor et al., 2013), given the emergency nature of their health condition(s). This approach prevents endogeneity problems associated with the patient self-selection into hospitals (Moscelli et al., 2021) with expected higher quality or lower changes in workforce composition, and also limits violations of the Stable Unit Treatment Value Assumption required to give a causal interpretation to our difference-in-difference estimates.

The core result of our empirical analysis is that hospital quality of care was negatively impacted by the change in the hospital staff composition, caused by the success of the referendum to leave the EU. After the Brexit referendum, for emergency patients admitted to a hospital with a mean pre-referendum share of EU nurses, the risk of in-hospital mortality increased by 5.31%, and the risk of unplanned emergency readmission increased by 2.28%. The estimate translates to about 4,454 additional hospital deaths in England during the post-referendum period and about 8,777 additional unplanned hospital readmissions.[3]

We show that hospital organizations substituted the missing EU nurses by hiring non-European nurses after the referendum, also because of the 2018 relaxation of the cap on healthcare workers' visas for non-EU migrants (Portes, 2022), so that no absolute nurse shortage arose. Most importantly, however, we show that the missing inflow of EU nurses, and the implied change in the hospital nursing workforce composition, is the main mechanism through which hospital care performance was affected. In particular, our analysis suggests that the NHS hospitals changed the workforce composition of new joiner nurses: newly hired nurses after the referendum were employed in lower salary bands, which can be interpreted as a sign of lower qualifications, experience and skills (Cortes and Pan, 2015). We also rationalize the decreasing quality of newly joining workers through a theoretical model of hospitals' optimal hiring rule. Our model predictions are consistent with the empirical patterns we document and provide a skill selection mechanism linking the Brexit referendum outcome to changes in hospital quality. Confirming our proposed mechanism, we further provide suggestive evidence that nurse satisfaction with the care quality they offered to patients fell after the referendum in the more exposed hospital organizations. Importantly, we also investigate and rule out alternative economic mechanisms. On the demand side, we show that the number of hospital patients, either prospective or treated, did not change with the intensity of the shock. On the supply side, we collect and put together publicly-available NHS hospital organization balance sheet data, showing that hospital expenditures and revenues from reimbursements of patient treatments were unaffected by the exposure to the migration shock. We ensure that the drop in hospital quality was not induced by capacity constraints, such as a lack of staff rather than a human capital loss. For example, we find no evidence of a productivity decrease in the hospital organizations more exposed to the shock, as the post-treatment share of occupied beds did not show any significant decrease.

Our work contributes to several lines of research in health, labour, organization and political economics. First, we provide evidence that public migration policies directly impact organizational performance. Most of the existent literature in this research area has investigated the effects of high-skilled migration policies and their impact on firm performance and innovation (Choudhury et al., 2022, Doran et al., 2022, Kerr and Lincoln, 2010, Peri et al., 2015, Hornung, 2014, Mitaritonna et al., 2017, Ottaviani et al., 2018, Terry et al., 2023). Only a little work exists on migrant workers' labour market effects in nursing markets (Furtado and Ortega, 2023, Grabowski et al., 2023), and to our knowledge none in hospitals. Our work, instead, contributes to literature on the effects of return migration (Adda et al., 2022, Borjas and Bratsberg, 1996, Dustmann and Görlach, 2016) and migration barriers (Lee et al., 2022), by originally investigating how the reverse migration policy shock due to the Brexit referendum vote induced the withdrawal of skilled migrants, as well as its effects on end-consumers, that is, hospital patients.

We contribute to the literature investigating the effects of staff turnover, and in particular the

[3]Similarly to increases in hospital-related mortality, unplanned emergency readmissions to hospital are a known indicator of poor hospital care quality (Gruber and Kleiner, 2012, Moscelli et al., 2018, 2021, Friedrich and Hackmann, 2021).

workers' substitutability, on firm performance. Other papers studying worker exit on firm or institutional performance either examine only short-run, transitional effects (Bertheau et al., 2022, Kuhn and Lizi, 2021), temporary absences such as parental-leave programs (Brenoe et al., 2024, Gallen, 2019, Ginja et al., 2023, Huebener et al., 2022), look at small entities (Becker et al., 2017, Brenoe et al., 2024, Gallen, 2019), or explicitly deal with productivity spillovers between workers (Jones and Olken, 2005, Huber et al., 2021, Waldinger, 2010, 2012). Instead, we exploit a large-scale employment shock that affected sizeable organizations – NHS hospital organizations – rather than using as a source of identification single exogenous layoffs or deaths in small groups of workers.

To our knowledge this is also the first paper to quantify one channel – that is, immigrant hospital workers – through which the outcome of the Brexit vote has causally impacted public health in the United Kingdom. This adds to the literature on the (unintended) economic effects of the recent deglobalization and nationalistic trends, and Brexit in particular (Born et al., 2019, Hantzsche et al., 2019, Fetzer and Wang, 2020, Davies and Studnicka, 2018, Breinlich et al., 2020, 2022). While these papers mainly focus on macroeconomic implications, we provide microeconomic evidence on healthcare services, with possible effects on all UK citizens and their daily lives. The effects of the Brexit referendum on public health are of particular interest from a political economy perspective, given that it has been found that regional heterogeneity in NHS performance was a driver of 'Leave' votes (Alabrese et al., 2019, Becker et al., 2017) and Brexit campaigners, including the former Prime Minister Boris Johnson, claimed that Brexit would have improved substantially the English NHS funding.

Furthermore, our work contributes to the literature that identifies the economic value of nurse and physician labour supply for patient health (among others: Fetzer et al. 2024, Foster and Lee 2015, Furtado and Ortega 2023, Gruber and Kleiner 2012, Chan Jr. and Chen 2023). Within the English NHS context, Propper and van Reenen (2010) exploit regional differences in the outside-option wage to control for unobservable factors affecting hospital nurse quits and nurse quality; whereas Friedrich and Hackmann (2021) exploit a parental leave program in Denmark, which led to a short-run decline in nursing, and find a mortality increase in retiree care homes. Rather than exploiting an absolute deficit of workers in some hospitals or care homes, our investigation makes use of an original identification setting, the sudden drop in the inflow of EU nurses after the Brexit referendum, which is a persistent shock that changed the hospital workforce composition.

Last, but not least, our study contributes to a number of studies documenting the effects of immigration in the United Kingdom (Dustmann et al., 2005, Manacorda et al., 2012, Dustmann and Frattini, 2014, Ottaviani et al., 2018).

The empirical and theoretical results of our work are informative for policymakers, and provide the following take-away message: prospective migrants, and especially high-skilled workers like nurses, are responsive to expected changes in immigration legislation and cultural hospitality of prospective host countries. Moreover, abrupt shocks to skilled workers' labour supply can affect an organization's productivity (in our case, the quality of hospital care provided). Therefore, countries whose labour markets rely on the inflow of foreign skilled

workers, such as the US, the UK and many other OECD member states, have to carefully weigh which labour market signal they relay to prospective migrant workers when more stringent immigration laws are proposed or approved.
The remainder of the paper is organized as follows. Section 5.2 describes the Brexit referendum history, the NHS workforce, and the data sources we used. Section 5.3 provides our theoretical framework. Section 5.4 presents the empirical strategy. Section 5.5 reports the main descriptive and estimation results. Section 5.6 investigates the mechanisms driving the main findings. Section 5.7 concludes.

5.2 Background

5.2.1 The Brexit Referendum

The Brexit referendum (BR) was announced on 20th February 2016 and took place on 23rd June 2016. On this date, British, Irish and Commonwealth adult citizens residing in the UK or Gibraltar were asked whether the UK should remain a member or leave the European Union. The BR was the culmination of a series of failed negotiations between the UK and the EU regarding the terms of the EU membership for the UK, especially with respect to policy matters like immigration and national sovereignty. The referendum had consultative nature, i.e. its outcome was not meant to be binding for the UK government. However, the British government of the time committed to implement the referendum result.
The BR had a 72.21% turnout, with 17,410,742 people (corresponding to 51.9% of the actual voters) voting in favour of leaving the EU. The official exit from the EU was dated 31 January 2020, but the UK remained a member of the European Single Market for a transition period lasting until the end of the year, which served to finalize the terms of the withdrawal and favour a smoother exit from the European Union. Starting from 1st January 2021, EU laws did not apply to the UK anymore, including the freedom of movement of persons and workers, which holds compulsorily within EU country members, as established by the 1992 Treaty of Maastricht, and allows any EU national to freely move and seek a job in another member state of the European Single Market (European Parliament, 2023). Thus, only from 1st January 2021 EU citizens willing to settle in the UK have been subject to the same migration rules of non-EU citizens and needed a visa to work in the UK. Between the BR date and 1st January 2021 there was no change in the immigration rules for EU workers moving to the UK, compared to the pre-BR regulation. Moreover, EU citizens already resident in the UK before the end of the transition period retained their pre-Brexit immigration rights under the so-called "EU settlement scheme", a dedicated scheme for EU nationals designated by the UK Home Office (House of Commons Library, 2020).

5.2.2 The NHS Workforce and Staff Recruitment From Abroad

The English NHS employs around 1.5 million people overall and it is one of the largest employers worldwide. It provides tax-funded, free at the point of use healthcare services to the general population, across more than 1,000 hospital sites grouped into 219 healthcare organizations called "Trusts". Nurses and doctors represent the core of the hospital workforce and account for more than one third of the total number of English NHS employees. As of January 2014 (2018), the English NHS employed 52,452 (59,253) hospital senior doctors and 338,333 (346,941) nursing and midwivery staff (NHS Digital, 2023*a*).

In order to work as a doctor or nurse for the English NHS, one must hold a relevant medical or nursing degree recognized by accredited bodies. Medical graduates who wish to become fully qualified doctors have to register with the General Medical Council (GMC) and undergo an in-hospital training programme to specialize in a given medical area. Junior medical workers account for approximately half of the total doctor workforce. Instead, nursing graduates can be hired immediately by English NHS hospitals as fully qualified nurses, upon registration with the Nursing and Midwifery Council (NMC).[4]

International medical and nursing graduates wishing to join the English NHS from abroad have to show respectively the GMC and the NMC that they possess a valid qualification and competence to practice. They also have to prove their knowledge of the English language. Foreign doctors' medical skills are evaluated by the so-called Professional and Linguistic Assessments Board (PLAB), consisting of a multiple choice test (PLAB1) and an objective structured clinical exam (OSCE) to be taken only in the UK. Similarly, foreign nurses' clinical skills get screened first through a computer-based test (CBT) and then by a practical OSCE competency test. Conditional on passing the PLAB1 (CBT), prospective doctors (nurses) can apply through one of the designated visa routes and enter the UK territory to take the second test. The successful completion of the OSCE exam provides the final clearance for joining the English NHS.

Starting from January 2021, the accreditation process described above applies to all international doctors and nurses, regardless of the foreign country in which they obtained their professional qualification. Instead, until the end of the Brexit transition period, EU laws automatically allowed doctors and nurses trained in a country of the European Economic Area (EEA) to practice as healtcare professionals in the English NHS, without the need to take either the PLAB or the CBT and OSCE exams. However, starting from January 2016 all nurses from the EEA willing to join the English NHS were required to present an English language proficiency certificate, which is a requirement in place since 2005 for international nurses trained in non-EEA countries. In June 2014, a similar language requirement for international doctors was extended to doctors trained in the EEA area.

5.2.3 Data Sources

We combined multiple data sources to create a unique patient- and nurse-level dataset which stretches from June 2012 to May 2019, namely the seven years around the Brexit referendum date. We abstract from years before this time window due to a major organizational re-

[4]Alternative, although less popular, routes to become a nurse in the English NHS are through the completion of a registered nurse degree apprenticeship (RNDA) directly offered by English NHS organizations or by joining the nurse workforce as a nursing associate.

structuring and merger wave among NHS organizations, which took place mainly until early 2012. Years after 2019 are dropped due to the onset of the COVID-19 pandemic in March 2020, which represented a major shock on the English NHS (Fetzer et al., 2024). Overall, our analysis sample comprises 131 acute care providers which consistently admitted patients over our period of study, thus excluding mental health providers or organisations undergoing any (potentially confounding) hospital consolidation event close to the Brexit referendum date.[5]

Electronic Staff Records. We use Electronic Staff Records (ESR) data, an administrative monthly worker-level payroll database, whose records on the universe of NHS hospital nurses and doctors include rich information on these clinical workers' demographics (e.g. age, nationality, sex) and employment-related variables, such as hours worked, earnings, staff grade and role, date of joining the NHS, and hospital organization of employment. We use the ESR data to compute inflows and outflows of NHS hospital nurses for different nationality groups, i.e. EU, British, non-EU nurses (and doctors), at the hospital organization level. To allocate nurses into nationality groups, we exploit the panel nature of the ESR data and rely on the first non-missing nationality record of each individual nurse. Based on their ethnicity background (e.g. White/Black British, White/Black Asian, White Irish, White European), we also impute the broad nationality group of almost all those residual nurses who never present a valid nationality information.[6] We define the hospital-level exposure to the Brexit shock based on the distribution of the shares of EU, British and non-EU foreign nurses of each hospital organization in the pre-referendum period. Moreover, we use the salary grades observed in the ESR records to track hiring decisions at hospital organization level before and after the Brexit referendum, and to proxy for nurse 'quality' – that we intend as a mixture of skills, experience and qualifications – when joining the NHS as new hires.

Hospital Episodes Statistics. We use Hospital Episodes Statistics (HES) Admitted Patient Care (APC) data to compute patient-level health outcomes that are frequently used as hospital quality indicators, such as the occurrence of death and unplanned readmissions, and other hospital-level indicators (e.g., the number of admitted patients). HES APC is an administrative database containing the universe of patient admissions to NHS acute care hospitals; its records provide rich information on patients' demographics, e.g. age, sex, date of admission, method of admission, medical conditions, income deprivation of the patient residence at small area level. We exploit this information to create binary indicators at patient level for in-hospital death and unplanned emergency readmissions within 30 days from the index hospital admission, as well as the Charlson index (Charlson et al., 1987), a known indicator of patient health risk due to pre-existing comorbidities. We also compute a binary indicator for patient mortality anywhere (in and outside the hospital) within 30 days from the index hospital admisison, by linking patient-level HES APC records to the Office for National Statistics (ONS) Civil Registration Deaths dataset, which holds information on the exact date of death of patients admitted to NHS hospital organizations.

[5]The information on mergers and acquisition events among NHS hospital organisations comes from the NHS Workforce Statistics Data Quality Annex regularly published by NHS England.

[6]In our sample, only 4.2% of the total number of nurses has a missing nationality record.

NHS Staff Surveys. The NHS Staff survey collects the self-reported assessments of NHS hospital workers with respect to several dimensions of their jobs. Its records consists of several hundred thousands NHS workers participating on an annual basis to this repeated survey, which is the largest longitudinal survey of a healthcare workforce in the world. We observe the occupation and the tenure of each worker matched with their answers. This allows us to identify the self-reported changes in quality of care and working environment conditions perceived by nurses employed in NHS hospital organizations with a different exposure to the Brexit referendum shock.

Hospital Trusts Financial Accounts. To test how the Brexit referendum affected the accounts and finances of NHS hospital organizations, we collected publicly available data on the annual financial reports of NHS hospital organizations.[7] This data includes the aggregate monetary £ pound value of healthcare expenditures related to operative hospital costs, which are mostly due to patient admissions.

Summary Statistics. The sample includes 131 acute care NHS hospital organizations with an overall number of 9.5 million emergency patients and 17.6 million emergency admissions in the pre-referendum period. *Panel a* of Appendix Table 5.B.1 provides the summary statistics on the health outcomes and covariates of the emergency patients in our sample.[8] The average in-hospital mortality risk, within 30 days from the index admission to hospital, was 3.3% in the pre-referendum period (June 2012 to May 2016) and 3% in the post-referendum period (June 2016 to May 2019). The risk of unplanned emergency readmission within 30 days from the index emergency hospital discharge, was 15.2% in the pre-referendum period and 16.1% in the post-referendum period. *Panel b* of Appendix Table 5.B.1 provides an overview of the clinical workforce composition of the hospital organizations. On average, in our pre-referendum sample, the acute care NHS hospital organizations employed approximately 1,661 nurses, out of which 98 (221) were foreign EU (non-EU) nurses, and 316 senior doctors, out of of which 27 (70) were foreign EU (non-EU) doctors. In the post-referendum period, the average acute care NHS hospital organization employed approximately 1,751 nurses, out of which 147 (246) were foreign EU (non-EU) nurses, and 359 senior doctors, out of of which 36 (88) were foreign EU (non-EU) doctors.

5.3 Conceptual Framework

We develop a simple conceptual framework to understand how the Brexit referendum might affect workforce composition and hospital quality in the NHS. This model also allows us to formalize and study the heterogeneous effects of Brexit across occupations and hospital organizations with different exposures to the shock.

Setup. Consider a hospital organization that wants to fill a mass $M < 1$ of vacancies.

[7] This data is available at `https://www.england.nhs.uk/financial-accounting-and-reporting/nhs-providers-tac-data-publications/`.

[8] The analogous information for all emergency and non-emergency hospital patients is reported in Appendix Table 5.B.2.

Suppose there is a unit mass of prospective workers, each deciding whether to apply for a job at the NHS. Each prospective worker might be from two possible origins $j \in \{e, r\}$, where e denotes "European Union" (EU nationals) and r the "rest of the world", with $\mu \in (0, 1)$ denoting the share of EU nationals. A worker i from origin j gets utility u_{ij} if they join the NHS and v_{ij} if they stay in their home country.[9]
The utility gain from joining the NHS is assumed to be

$$u_{ij} - v_{ij} = \omega_j + \gamma_j - \varepsilon_i,$$

where $\omega_j \in \mathbb{R}$ denotes the average expected present value wage gain and $\gamma_j \in \mathbb{R}$ other expected non-wage benefits of joining the NHS when coming from region j, while $\varepsilon_i \in \mathbb{R}$ denotes a mean-zero idiosyncratic preference shock.[10]
We assume that each worker i has a skill level $\theta_i \in \mathbb{R}$ that affects the quality of care as defined later. We impose that $(\theta_i, \varepsilon_i)$ are independent of each other, i.i.d. across workers, admit a continuous probability density function, and have finite first moments. We denote by F (by f) the cumulative distribution (probability density) function of ε_i and by G (by g) the one of θ_i. To stress that our results are not given by exogenous differences across worker groups, we assume the same skill and preference shock distributions irrespective of the worker origin.[11]
If hired, each worker's type provides a quality of care $q(\theta)$, where $q : \mathbb{R} \to \mathbb{R}$ is strictly increasing. The hospital organization then decides what share of workers of each type to hire. The total quality of care provided by the newly hired workers is

$$Q(h, a) = \int q(\theta)h(\theta)a(\theta)d\theta,$$

where $h : \mathbb{R} \to [0, 1]$ denotes the share of the applicants of a given type the hospital hires, and $a : \mathbb{R} \to \mathbb{R}_+$ denotes the number of applicants of each type θ.
A potential worker of origin j applies to the NHS if $u_{ij} - v_{ij} \geq 0$, or equivalently if $\varepsilon_i \leq \omega_j + \gamma_j$. Therefore, the mass of applicants of each type is

$$a(\theta) = \Big[\mu F\big(\omega_e + \gamma_e\big) + (1 - \mu)F\big(\omega_r + \gamma_r\big)\Big]g(\theta).$$

We assume the total mass of applicants would be sufficient to cover vacancies if all were accepted, that is, $M < \int a(\theta)d\theta$.

The Hospital's Problem. The hospital observes the set of applicants and wants to maximize the total quality of care subject to the constraint that they can hire at most a mass M

[9] The insights of the model hold regardless whether the r group includes British workers. For simplicity, we keep the main analysis with only two groups: EU and non-EU nationals.

[10] The non-wage benefits γ_j may include monetary and non-monetary benefits, such as access to a pension scheme, stability at work, gains from moving to the UK, or an intrinsic value for the job.

[11] The results remain qualitatively unchanged if different origins are associated with different skill or preference shock distributions.

of workers. The hospital solves

$$\max_{h:\mathbb{R}\to[0,1]} \int q(\theta)h(\theta)a(\theta)d\theta. \tag{5.1}$$

subject to

$$\int h(\theta)a(\theta)d\theta \leq M. \tag{5.2}$$

Lemma 1. *The hospital accepts all applicants with type above the cutoff* $\theta^* := max\{\theta_0, \tilde{\theta}\}$*, where* θ_0 *and* $\tilde{\theta}$ *are defined as*

$$\theta_0 := inf\{\theta \in \mathbb{R} : q(\theta) \geq 0\} \text{ and } M = \int_{\tilde{\theta}}^{+\infty} a(\theta)d\theta.$$

Lemma 1 shows that the hospital hires the most skilled workers until they fill all their vacancies or reach a minimum acceptable skill level. Any worker with skill $\theta < \theta_0$ negatively impacts the quality of care. Hence, we refer to them as unqualified for the job. If there are enough qualified applicants, the hospital uses all its budget for new hires, and the hiring skill cutoff is $\tilde{\theta}$. Otherwise, the hospital hires all applicants with skills above θ_0 but fails to fill all the vacancies.

Brexit Referendum. We model the Brexit referendum effects as a decrease in the EU nationals' future discounted expected payoff from moving to the UK. This decrease stems from potential EU national movers' revised expectations about direct future monetary and non-monetary losses the Brexit enforcement regulation might cause to EU national workers based in the UK. For example, they include increased costs for travels, visa, recognition of overseas-acquired qualifications, settlement hurdles as EU-migrants to the UK (UK Government, 2020), as well as the immediate disutility related to an increase in the anxiety and uncertainty about the future (Frost, 2020, Teodorowski et al., 2021) in terms of employment-related, political and civil rights.[12] Formally, we say that $\gamma_e^{pre} > \gamma_e^{post}$. We then denote by θ^*_{pre} and θ^*_{post} the hiring skill cutoff pre and post-referendum and study how the BR affected hospitals' hiring cutoff, quality of care, and prevalence of worker shortages.

Proposition 1. *Suppose that* $\theta^*_{pre} > \theta_0$ *and* $\gamma_e^{pre} > \gamma_e^{post}$*. Then, in the post-referendum*

1. *the hiring skill cutoff decreases;*
2. *the EU-worker joining rate decreases;*
3. *the quality of care decreases;*
4. *worker shortages do not occur unless* $(\gamma_e^{pre} - \gamma_e^{post})$ *is sufficiently high.*

Proposition 1 delivers several insights about the referendum's effects on the workforce composition and the quality of care. First, it shows that the decrease in the non-wage gains of joining the NHS for EU nationals reduces the overall supply of workers, and to fill all their vacancies a NHS hospital organization needs to decrease its hiring standards. Second, it shows

[12] See also KPMG (2017), Nursing Times (2018), The Guardian (2019) and Financial Times (2019).

that simultaneously with a decrease in the hiring standard, one also observes a decrease in the share of EU workers and a decrease in the overall quality of care. Finally, Proposition 1 shows that a decrease in the quality of care occurs even when there is no increase in worker shortages: a NHS hospital organization might be able to fill all its vacancies, yet the decrease in the UK attractiveness to EU nationals harms the selection of skilled workers regardless of their country of origin and thus reduces quality of hospital care.

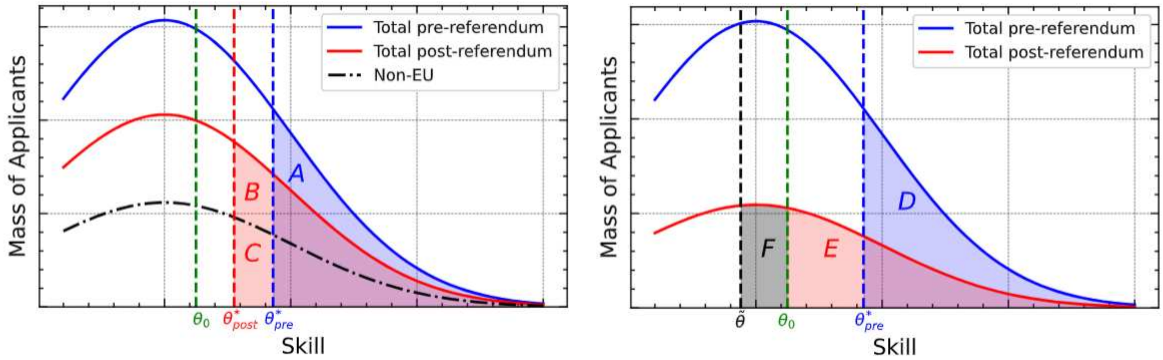
It is important to note that the resulting hiring skill cutoff θ^* is the same for EU and non-EU workers pre and post-referendum. The decrease in the quality of care stems from substituting higher-skill EU workers with lower-skill workers of any origin, i.e. Europeans, non-Europeans, and British. Figures 5.1 and 5.2 illustrate this substitution pattern when we observe worker shortages and when we do not, as well as across low and high-exposure hospitals (low vs. high μ).

Figure 5.1a describes the effects of a reduction on γ_e that does not cause an increase in worker shortages. The solid blue (red) line denotes the pre(post)-referendum total number of applicants with a given skill level, while the dash-dotted black line plots the number of non-EU applicants. The difference between the solid blue (red) and dash-dotted black line denotes the number of EU applicants pre(post)-referendum. The area shaded in blue (red) represents all the hired workers pre(post)-referendum. Area A — the area in-between the blue and red solid lines — denotes the mass of higher-skill workers who would apply prior to the referendum but do not after it, while B and C display the mass of workers who are hired after the referendum but would have not been hired if the hiring skill cutoff had not decreased. When comparing the workforce composition pre and post-referendum, there is a substitution from higher-skill workers (area A) to lower-skill workers (area $B + C$), which reduces the quality of hospital care. Moreover, note that the higher-skill workers who no longer apply (area A) are all EU nationals, while the new hires below the pre-referendum hiring skill cutoff are from both EU (area B) and non-EU (i.e., British and non-EU nationals; area C) countries of origin. Consequently, the quality of care and the share of new EU joiners simultaneously decrease.

Figure 5.1b displays the changes in the workforce composition of a reduction in γ_e that instead causes also worker shortages. When the decrease in the attractiveness of the NHS for one group of prospective workers is big enough, the NHS hospital organizations are unable to find a sufficiently large number of qualified workers (with $\theta \geq \theta_0$) to fill all of their vacancies. The hiring cutoff then becomes the minimum qualification standard θ_0. Area D represents the mass of higher-skill workers that stopped applying after the decrease in γ_e, area E denotes the newly hired workers that would not have been hired absent the reduction in the hiring skill cutoff, and area F is the size of the shortage, meaning the mass of vacancies that remain unfilled.

The next result compares how different hospitals are affected by the same shift in γ_e. Proposition 2 shows that hospitals that have a larger share of EU national potential applicants (higher μ) are more affected by the referendum's result, as long as the average utility gains from moving to the UK to work for the NHS are larger for workers coming from the rest of

Figure 5.1: Model implied applicant pool and hiring cutoffs.



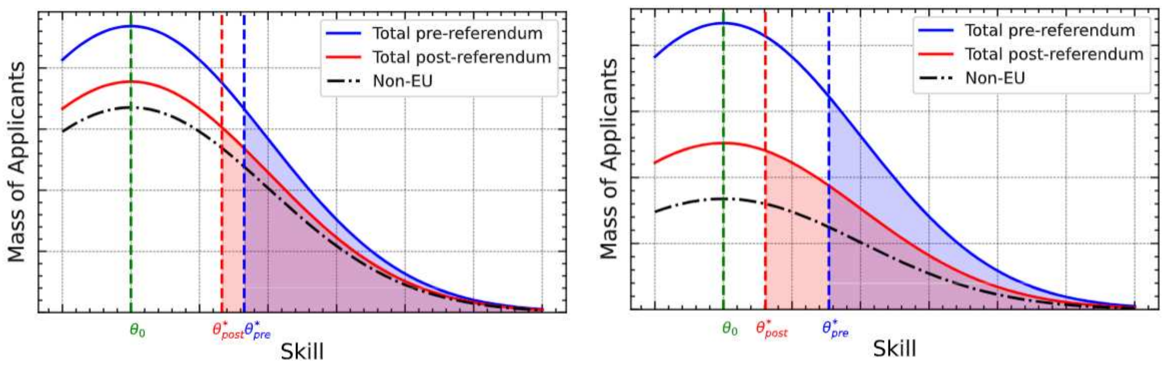
(a) No worker shortages (b) Increased worker shortages

Note: This figure gives the effects of a change in γ_e on the workforce composition under no and increased worker shortage.

the world than from EU countries. Figures 5.2a and 5.2b illustrate the result by, respectively, plotting the effects in a low-exposure hospital (low μ) and a high-exposure hospital (high μ). The shaded area in red between θ^*_{pre} and θ^*_{post} is much smaller in the low-exposure Figure 5.2a than in 5.2b. The intuition is straightforward: in high-exposure hospitals, the decrease in the supply of workers is larger, and the hospital must hire more workers with skills below its original hiring skill cutoff.

Proposition 2. *Suppose the average utility gain from joining the NHS is larger for workers from the rest of the world than from EU countries ($\omega_r + \gamma_r > \omega_e + \gamma_e$). Then, the number of unfilled vacancies plus the number of workers hired after the referendum with skills below the pre-referendum cutoff increases in the share of potential EU national applicants μ.*

Figure 5.2: High versus Low Exposure.



(a) Low-exposure Hospitals (b) High-exposure Hospitals

Note: This figure gives the effects of a change in γ_e on the workforce composition for hospitals of different exposure (measured in μ).

Finally, we show that the effect of a decrease in γ_e on the set of hired workers vanishes as the *absolute* wage gains for the job increase. That is, as the ω's increase, the difference between θ^*_{pre} and θ^*_{post} goes to zero. The intuition is that the decrease in non-wage gains of moving to the UK may be similar across occupations, while the absolute monetary gains vary with each occupation's pay level. Even if all occupations receive a similar relative wage gain, meaning the same percentage increase compared to what they were paid in their country of origin, the absolute gain is much larger for the ones in high-paying jobs. This result suggests that we should observe stronger effects of the referendum for relatively lower-paid occupations, such as nurses, compared to better-paid ones, such as medical doctors.

Proposition 3. *The difference between pre and post-referendum hiring skill cutoff goes to zero as the wage gains to joining the NHS increase.*

As a quick summary of the main takeaways of our conceptual framework:

1. The quality of care decreases after the referendum, even when there are no additional worker shortages or a decrease in the total number of workers.
2. The decrease in the quality of care is driven by a decline in the skill level of newly hired workers, irrespectively of their country of origin.
3. Hospitals that more heavily relied on EU workers hire a larger number of workers below the pre-referendum skill cutoff.
4. The referendum effects vanish for sufficiently highly paid occupations.

5.4 Empirical Strategy

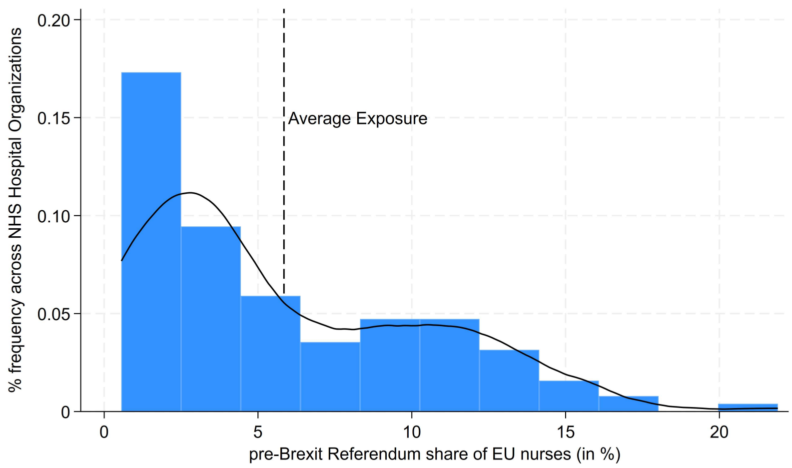
Our empirical strategy consists of three steps. In the first two steps, we evaluate the effects of the Brexit referendum on the clinical workforce composition of NHS hospitals. In particular, the first step provides descriptive evidence about any structural break in the workforce composition of both nurses and senior doctors employed in all NHS acute hospitals. The second step provides regression-based evidence on the effects of the Brexit referendum shock on nursing workforce composition only – as we show that nurses are the occupational group predominantly affected by the outcome of the Brexit vote – with respect to the final sample of 131 acute care NHS hospitals for which we measure both patient health outcomes and hospital workforce composition. Building on the previous two descriptive steps, we then estimate the causal effect of workforce composition on hospital quality (third step).

In our third step of the empirical strategy, we exploit the insight that hospital organizations that relied more heavily on EU nurses in the pre-referendum period were more exposed to the missing inflow of EU nurses after the referendum. This allows us to study whether the Brexit-induced net loss of EU nurses affected English hospital performances, as we are interested in examining the effects of the employment shock induced by the Brexit referendum on hospital care quality. To do so, we rely on the following event-study design with a continuous treatment and estimated at the patient level:

$$Y_{i,h,t} = \alpha_h + \lambda_{r,t} + \sum_{\substack{k=2012/2,\\ k\neq 2016/1}}^{2019/1} \beta_k(\mathbb{1}[t=k]_t \times \overline{EU}_h) + X_i'\theta + \epsilon_{i,h,t} \tag{5.3}$$

$Y_{i,h,t}$ is patient i's health outcome following admission to NHS hospital h on date t; for example, $Y_{i,h,t}$ is an indicator variable valued one if patient i died within 30 days from her admission to hospital h in month t, and zero otherwise. $\overline{EU}_h$ is our treatment exposure, the share of European nurses in hospital organization h, averaged over the four years preceding the referendum date, i.e. from June 2012 to May 2016.[13] This historical average approximates the exposure of each hospital organization to the effects of the Brexit shock, summarized in subsection 5.5.1. The distribution of this treatment variable at the hospital organization level is plotted in Figure 5.3. The 'average' hospital organization has an exposure of 5.84%, that is, about six out of 100 nurses employed are from the EU.

Figure 5.3: Distribution of the treatment exposure



Note: This figure gives the hospital organization-level treatment exposure, i.e., the average share of EU nurses in the pre-referendum period of our analysis window. Mean = 5.84%; Standard Deviation = 4.58%; Minimum = 0.56%; Maximum = 21.90%.

We interact this exposure variable with half-year bins, namely dummy variables that take the value of one in each of the six-month periods around the referendum date. Indeed, because the BR took place in June 2016, we define relative half-year indicators based on the time windows that range from June to November and December to May. We estimate Equation 5.3 by choosing the interaction of the treatment exposure with the relative semester right before the referendum as the reference (omitted) category.[14] α_h and $\lambda_{r,t}$ are hospital and NHS region-specific half-year fixed effects, which capture time-invariant, hospital-specific quality differences as well as regional shocks over time. The NHS region fixed effects allow for a more granular comparison of hospitals within the same region, making our approach less sensitive to heterogeneous economic reactions to the Brexit referendum outcome in space.

X_i is a rich vector of patient-level characteristics that are likely correlated with the patient outcome variable. The vector X_i includes: admission-method indicators (e.g. admission via

[13] Specifically, our treatment exposure measure has been computed as the four-year average of the monthly hospital share of EU nurses: $\overline{EU}_h = \frac{\sum_{m=2012/06}^{2016/05} EU_{h,m}}{48}$, where $EU_{h,m}$ denotes the share of European nurses in hospital organization h and calendar month m.

[14] For the outcome variables on which we only observe annual data, year 2015 is the last pre-treatment time period, and as such used as the reference event-study period in the estimation.

Accident & Emergency services or via general practictioner), 19 diagnosis indicators based on the ICD-10 classification, a female indicator, seven age-bracket indicators (18-29, 39-39, 40-49, 50-59, 60-69, 70-79, 80+) and income deprivation indicators (one for each quintile of the national distribution). We also control for seasonal patterns in mortality risk through the inclusion of admission-month indicators, and for the observed patients' frailty with a polynomial of degree two in the Charlson Comorbidity Index (CCI). These covariates are standard controls in models predicting hospital patients' outcomes (Cooper et al., 2011, Gaynor et al., 2013, Moscelli et al., 2021), as they serve to account for the heterogeneity in patient-level case-mix across different hospitals (NHS Digital 2023*b*).

Identification Assumptions and Threats. The coefficients β_k measure how patient- and hospital organization-level outcomes evolved relative to right before the referendum date, for different levels of hospital exposure to the BR shock. More specifically, the post-referendum interaction terms indicate whether hospital organizations that were historically more reliant on EU nurses experienced, for example, a post-referendum increase in the average patient's mortality risk, relative to hospital organizations that were less reliant on EU nurses.

Two main assumptions have to be satisfied in order for Equation 5.3 to estimate a causal effect in the proposed framework. First, the 'parallel trends assumption' (PTA) requires that patient care quality in hospitals of heterogeneous exposure to the shock would have developed similarly had the referendum not occurred. A visual inspection of the pre-referendum interaction terms will be informative about whether such diverging trends existed or not. The absence of pre-trends will also help to rule out potential reverse causality concerns, namely the possibility that changes in hospital quality affected the pre-referendum pool of international workers and, by that, the hospital treatment status. In the presence of reverse causality, the evolution of patient outcomes over time would considerably differ between more and less exposed hospitals, prior to the Brexit referendum and, hence, would be visible in the form of pre-trends.

Second, Equation 5.3 requires no treatment spillovers across hospitals of different exposure levels, also known as the 'stable unit treatment variable assumption' (SUTVA). In our setting, treatment spillovers among hospitals in the vicinity could arise. Patients could, for example, switch hospitals in response to worsening health care in a highly exposed hospital, effectively smoothing treatment across hospitals. This instance, however, would imply that Equation 5.3 will underestimate most of our treatment effects of interest in absolute terms. Also, we explicitly test for such patients' behavioral response through a series of robustness checks. For example, in Table 5.B.3, we test for changes in patient composition as well as the catchment area population across hospitals of different exposures, documenting no effects in terms of key demographics and the overall number of admitted patients. Following the Brexit referendum, more exposed hospitals only seemed to have admitted less fragile patients, a potential source of downward bias which we control for through the inclusion of the comorbidity polynomial introduced above.

Moreover, our main estimation sample consists of only patients admitted to hospital for an emergency condition. Emergency patients have little to no choice over the hospital in which they are admitted, as by NHS clinical guidelines they need to be taken to the nearest hospital with capacity (Gaynor et al. 2013). Moreover, in England only NHS hospitals provide emergency care services to patients, so there is no outside option for emergency patients to be treated by private hospital providers. Given these institutional features, examining the health outcomes of emergency patients has the great advantage to limit the potential violations of SUTVA that would arise from the strategic behaviour of patients with respect to changes in hospital workforce composition and quality of care.[15]

Another possible concern is the confounding by time-varying unobservable factors that operate at hospital organization level and are correlated to, but not caused by, the labour supply shock which took place in the exposed hospitals after the Brexit referendum. To address this issue, we test a series of alternative mechanisms, such as the effects of the Brexit shock on hospital revenues, expenditures and bed occupancy rates.

Lastly, recent advances on difference-in-differences (DiD) show that DiD models with continuous treatment require a more demanding form of the parallel trends assumption (Callaway et al., 2024). For this reason, we will provide estimation results not only with our continuous measure of exposure, $\overline{EU}_h$, but also with a binary treatment indicator for hospitals that have an exposure value above the 75th percentile. This specification will identify how the quality of patient care changed after the Brexit referendum in highly exposed hospitals, relative to hospitals that belong to the first three quartiles of the exposure distribution.

5.5 Results

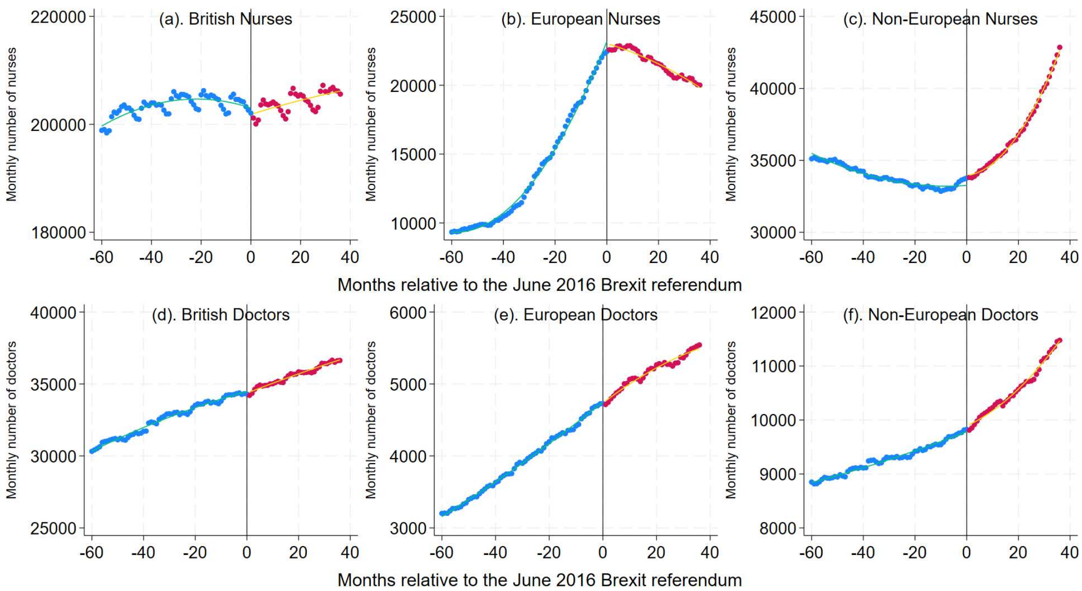
5.5.1 Brexit Referendum Shock and NHS Hospital Staff Composition: Descriptive Evidence at NHS Level

In this section we examine the effects of the June 2016 Brexit referendum on the nursing and medical workforce of the entire English NHS acute care hospital sector. Figure 5.4 provides a breakdown by nationality groups of the staffing levels of nurses and senior doctors employed in acute care NHS hospitals. Following the Brexit referendum, the total number of EU nurses in the English NHS started to fall (*panel b*), while that of non-EU nurses started to increase sharply (*panel c*). Instead, the total number of British nurses remained roughly constant, especially in the short-term (*panel a*). Similarly, there was no substantial discontinuity in the number of doctors of any nationality group around June 2016 (*panels d, e* and *f*). These patterns are consistent with our conceptual framework, where we have shown that the referendum effects on new joiners should vanish for better-paid occupations (e.g. doctors) compared to relatively lower-paid ones (e.g. nurses).

Figure 5.5 shows that the decrease in the overall number of European nurses and the increase in the overall number of non-European foreign nurses respectively stemmed from a marked decrease in the number of monthly joiner EU nurses (*panel a*) and a steady increase in the

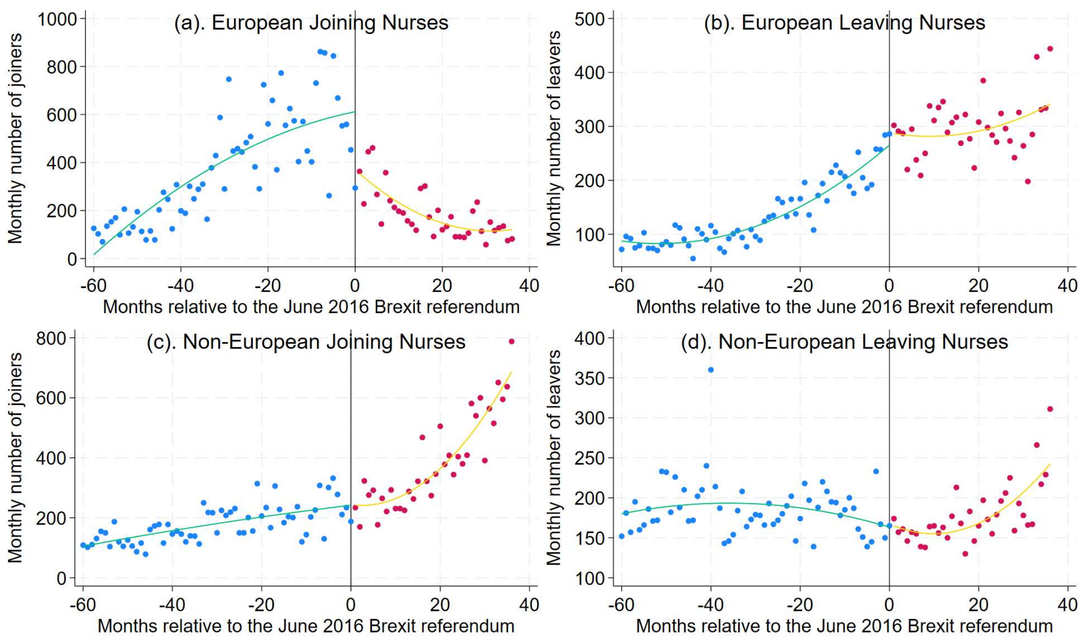
[15] As we show in Sections section 5.5 and section 5.6, we find no evidence supporting these concerns; moreover, the estimates of interest computed on the sample of both emergency and non-emergency patients are qualitatively very similar to those on emergency patients only.

Figure 5.4: Number of nurses and doctors employed in NHS hospitals, by nationality group



Note: This figure shows the number of nurses and doctors employed in NHS hospitals over time with the Brexit referendum date (June 2016) in month 0. Nurses' and senior doctors' nationality is classified according to the first non-missing nationality record (if present). The EU group includes Iceland, Norway, and Switzerland, namely all countries that have access to the European Single Market although not being formal EU member states.

Figure 5.5: Nurses joining and leaving NHS hospitals, by nationality group



Note: This figure gives the number of NHS monthly joiners and leavers by nationality group with the Brexit referendum date (June 2016) in month 0.

monthly joiner non-EU foreign nurses (*panel c*). The leaving rates of both EU and non-EU foreign nurses also increased (*panels b* and *d*), but less than the absolute changes in the number of joiners. The EU nationality workers already employed and settled in the UK prior to the BR would have made personal and professional investments, such as fostering their career within NHS hospitals, marrying, forming a household, having children; any decision to abruptly relocate outside the UK would have implied substantial divestment costs. Hence, EU nurses employed in the English NHS before the BR were likely less sensitive to the BR outcome and left the NHS at a slower pace than prospective EU nurses who choose not to move to the UK and be employed by the NHS. Moreover, Figure 5.B.1 in the Appendix shows that the BR resulted in large decreases in the joining rates of EU non-registered nurses (*pay bands 1-4, panel a*) and EU newly registered nurses (*pay band 5, panel b*), and a smaller decrease in the European senior nurse joiners (*pay bands 6-9, panel c*).[16]

To summarise, the outcome of the Brexit referendum led to a substitution between European and non-European nurses in the English NHS, which was almost entirely driven by a lower number of new nurses joining the healthcare system.[17] [18]

5.5.2 Brexit Referendum Shock and NHS Hospital Staff Composition: Hospital-level Regression-based Evidence

In this section, we show how the Brexit referendum shock impacted the workforce composition of our final analysis sample of 131 acute NHS hospitals at the hospital organization instead of the aggregate level. Before 2016, NHS hospitals differed largely in their reliance on EU nurses: before the referendum in 2016, the average share of EU nurses was 5.84%, with a range between 0.56% and 21.90% and a standard deviation of 4.58% (see Figure 5.3). Hospitals with a larger share of EU nurses before the referendum were more exposed to the negative labor supply shock induced by the Brexit referendum, due to EU nurses either reducing their migration inflow to NHS hospitals from EU countries as new joiners, or leaving their employing NHS hospital and the UK.

In Table 5.1, we report the association between the share of EU nurses (with respect to all nurses employed in the NHS hospital organization) during the four pre-referendum years (our continuous treatment exposure) and the hospital-level changes in key employment variables

[16]This finding chimes with evidence that the Brexit referendum had a disproportionate impact on lower-skilled and essential migrant workers (Sumption and Fernandez Reino, 2018, Fernández-Reino and Kierans, 2020).

[17]Figure 5.B.2 reports the corresponding graphs for English NHS doctors. Consistently with Figure 5.4, around the Brexit referendum date there was no considerable change in the joining and leaving rates of both European and non-European doctors.

[18]We have assumed that the drop in EU nurses joining numbers was entirely due to the uncertainty triggered by the BR outcome. However, the new English language requirement imposed since January 2016 on prospective nurses from the EU (see Section 5.2.2) could have acted as a concomitant driver of such drop. To assess which of these two shocks was mostly responsible for the aforementioned drop, in Figure 5.B.4 we plot the yearly percentage-point change in the monthly number of EU joiner nurses around the January and June 2016 dates. Figure 5.B.4 shows that the sharp reduction in the number of joiner nurses from Europe started only since June 2016, and also that the introduction of the new English language requirement had little to no impact on the EU nurse joining rate. A similar case of relative irrelevance occurred for EU hospital doctors, whose NHS joining rate did not change when a similar English language requirement was introduced in June 2014 (as shown by *panel a* of Figure 5.B.2).

between the pre- and post-Brexit period. *Panel a* of Table 5.1 reports the association between the hospital exposure to the Brexit referendum and the change in the share of nurses by nationality groups, which is computed as the change in shares between the two endpoints of our pre- and post-referendum periods (May 2019 and May 2016, respectively), in order to capture the full extent of the effects of the Brexit referendum shock on the NHS hospital nursing workforce composition, as displayed in Figure 5.4. Consistently with our hypothesis, hospital organizations that relied on a higher share of EU nurses before the referendum experienced a stronger decline in the share of EU nurses and a stronger growth in the number of non-EU nurses, after the Brexit referendum; the share of British nurses was unaffected, and there was no effect on total employment levels, as shown in Appendix Table 5.B.4; these findings are robust across and within NHS health regions.

Table 5.1: Exposure to Brexit shock and changes in nurse employment shares

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Employed nurses	Δ Share of British nurses		Δ Share of EU nurses		Δ Share of Non-EU nurses	
Pre-BR share of EU nurses	-0.1066*	0.0222	-0.2196***	-0.3840***	0.3290***	0.3586**
	(0.0558)	(0.1232)	(0.0588)	(0.1164)	(0.0775)	(0.1417)
Panel B: NHS joining nurses	Δ Share of British joiners		Δ Share of EU joiners		Δ Share of Non-EU joiners	
Pre-BR share of EU nurses	-0.1373	0.7538*	-1.0850***	-1.9434***	1.2087***	1.1746**
	(0.2101)	(0.4485)	(0.2491)	(0.4566)	(0.26410)	(0.4966)
Panel C: NHS leaving nurses	Δ Share of British leavers		Δ Share of EU leavers		Δ Share of Non-EU leavers	
Pre-BR share of EU nurses	-0.0143	-0.0827	0.3727***	-0.0971	-0.3314***	0.1857
	(0.1386)	(0.3452)	(0.1060)	(0.2520)	(0.1067)	(0.2062)
NHS region FE	No	Yes	No	Yes	No	Yes

Note: N = 131 acute care NHS hospital organizations. The changes in employment shares used as dependent variables in *panel A* is computed between May 2019 and May 2016, namely the two endpoints of our post- and pre-BR analysis periods. The changes in the total number of joining and leaving nurses used as dependent variables in *panels B* and *C* are computed as the percentage point changes in the cumulative number of joining or leaving nurses between the whole post-BR and pre-BR periods. Robust standard errors reported in parenthesis. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

According to the estimates in Table 5.1 *panel b*, NHS hospital organizations substituted the 'lost' inflow of EU nurses with non-EU nurses; such substitution is particularly evident among newly hired nurses: NHS hospitals with higher pre-referendum exposure to EU nurses hired a lower share of EU nurses but a higher share of non-EU nurses, cumulatively after the referendum. Moreover, estimates in Table 5.1 *panel c* show that there is no association between the referendum exposure and the change in the share of NHS leavers by nationality subgroup, confirming that the changes in workforce composition occurred predominantly through a substitution between EU and non-EU joiner nurses.[19] In Appendix Table 5.B.5, we also show that the pre-treatment exposure is also associated with a loss of South European and Irish nurses and an increase in Asian nurses, primarily driven by the reduced share of South European joiners and increasing share of Asian joiners (see Appendix Table 5.B.6).

[19]To compute the changes in leavers used as outcomes in *panel b* of Table 5.1, we use the cumulative number of leavers in the whole pre- and post-referendum periods.

5.5.3 Effects on Hospital Quality of Care

In this section, we present the results of our analysis, that is, the causal effect of changes in hospital workforce composition on a number of hospital care quality outcomes.

Baseline Results. Figure 5.6 presents the event-study estimates of the impact on the health outcomes of patients admitted for an emergency condition. Emergency hospital patients represent the marginal patients with the greatest health risk, for which changes in the workforce quality and composition could be more impactful. Our event-study regressions reveal that patients admitted to hospitals with a higher exposure to the Brexit referendum shock experienced higher risks of in-hospital mortality, mortality anywhere (in- and out-of-hospital) and also unplanned emergency readmission, in the post-referendum period. These effects are persistent over time, indicating that the decrease in healthcare quality was not just transitory, and the lead effects in the pre-referendum period are not significant. The mortality effect is robust, regardless of the use of a continuous or binary treatment exposure variable (which takes the value of one if a hospital organization is in the top quartile of the exposure distribution), whereas the readmission effect is more precise using a continuous treatment exposure variable. As shown in Appendix Figure 5.B.5 and Appendix Table 5.B.8, we find very similar effects using the sample of hospital patients that includes also planned (non-emergency) patients.

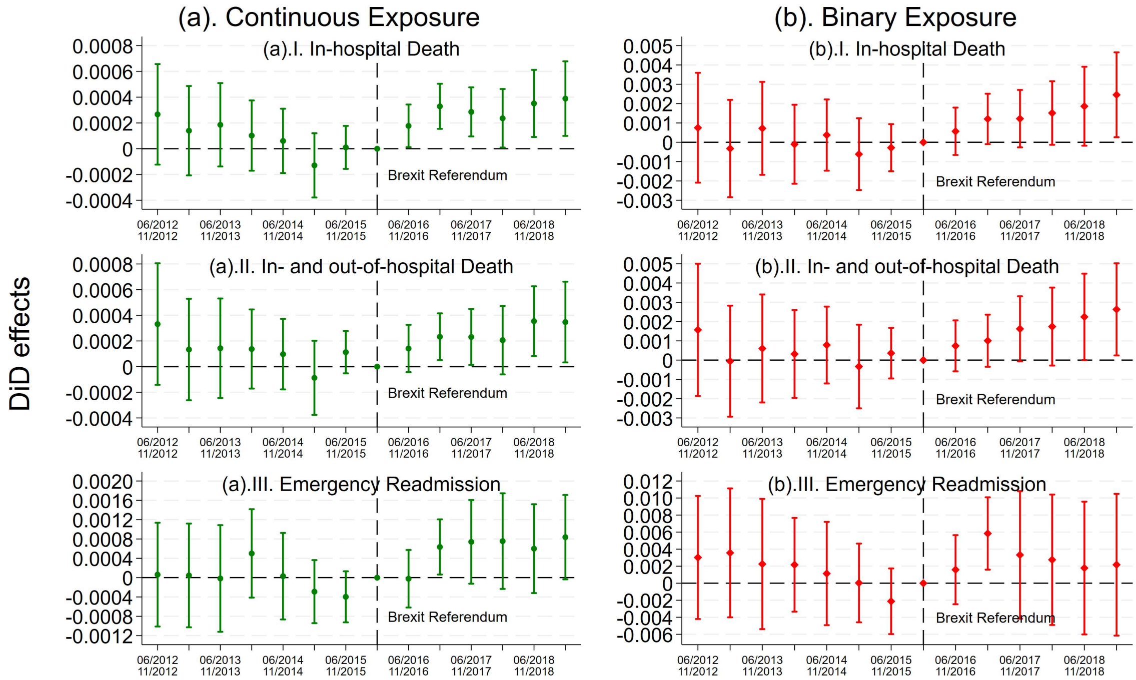
Implied Magnitude of the Effects. For the event-study based on the continuous exposure measure, the left axis of the event study plot provides the effects of a 1 percentage point (p.p.) increase in the Brexit shock exposure on the in-hospital mortality risk (within 30 days from hospital admission). Thus, for a hospital with mean exposure to the treatment (equal to 5.84%, see Figure 5.3), the in-hospital mortality risk increased by 5.31%, in- and out-of-hospital mortality risk increased by 3.45%, and the unplanned emergency readmission risk increased by 2.28%.[20]

These effects on patient health outcomes are equivalent, over the entire post-BR period, to 34 additional in-hospital deaths and 67 additional readmissions in hospitals with a mean pre-BR exposure measure (5.84%), for a total of 4,454 additional deaths and 8,777 additional readmissions.[21] To put things into perspective, the yearly magnitude of our estimated effects of the Brexit shock on in-hospital deaths (about 1,485 additional deaths per annum) is about half of the impact of the COVID-19 pandemic on non-COVID-related excess mortality (3,050 deaths) throughout the first twelve months of the COVID pandemic, as estimated by Fetzer et al. (2024). Given a decline in the number of EU nurses employed by the NHS acute care

[20] Each value is obtained as the product of the mean pre-referendum exposure variable (5.84) and the average of the post-referendum semestral effects on patient health outcomes using the continuous treatment event-study specifications (columns 1-3, Table 5.B.7), divided by the average pre-referendum health outcome risk (as reported in Table 5.B.1). Thus, respectively: $0.00029667 \times 5.84/3.26\%$ (in-hospital mortality); $0.00025167 \times 5.84/4.26\%$ (in- and out-of-hospital mortality); $0.00059333 \times 5.84/15.2\%$ (unplanned emergency readmissions).

[21] Our estimates have been obtained by multiplying the average post-treatment effects displayed in Table 5.B.7 (respectively equal to 0.00029667 for in-hospital mortality and 0.00059333 for unplanned emergency readmissions) by 113,247.5, which is the average number of patient emergency hospital admissions in the post-referendum period for the 131 hospital organizations in our sample, based on a total of 14,835,419 post-referendum patient emergency hospital admissions.

Figure 5.6: Effects of Brexit exposure on Patient Health Outcomes (emergency patients)



Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating Equation 5.3 with the continuous and the binary (i.e., below/above 75th percentile exposure) exposure. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals. The joint Wald test F-stat values, and their related p-values, testing the null hypothesis that all the leads effects are equal to zero, are reported in Appendix Table 5.B.7.

hospitals in our sample equal to 2,047 workers within the first three years after the referendum (see also Figure 5.4), the 'loss' of each EU nurse is associated with approximately 2.18 extra in-hospital patient deaths during our post-referendum period.[22] Considering an average age of 78.85 years for the hospital patients dying in our sample, an average life expectancy of 81.26 years in the UK in 2018 according to the World Bank, and assuming that the value of one year in full health is worth £60,000 (Glover and Henderson, 2010, Cutler and McClellan, 2001, Ryen and Svensson, 2015), the implied monetary life value 'lost' due to the Brexit shock was worth £644,048,400 (that is: $£60,000 \times 4,454 \times (81.26 - 78.85)$).[23]
Similarly, considering an average expenditure of £2,100 per each unplanned hospital readmission (Billings et al., 2012), the implied cost of the Brexit shock translates into additional £18,431,700 (that is: $£2,100 \times 8,777$) for NHS hospitals, or £140,700 for each of the 131 NHS hospital organizations in our analysis. Therefore, based on the 2016 regulated NHS nurse salary scales, the NHS hospital sector could have afforded hiring an additional number of either 841 (Pay Band 5) nurses, or 700 (Pay Band 6) senior nurses, or 587 (Pay Band 7) advanced nurses, if the £18.4 millions extra costs for the unplanned hospital readmissions

[22] In March 2016 (2019), there were 19,720 (17,643) EU nurses employed across the NHS hospital organizations in our sample, with an average of 150.53 (134.91) EU nurses employed per NHS hospital organization.
[23] For the life expectancy in the UK, see the World Bank data at: `https://datacommons.org/tools/timeline#&place=country/GBR&statsVar=LifeExpectancy_Person`.

caused by the Brexit referendum shock was averted.[24]

Heterogeneity by Patient Characteristics. We investigate how the effects on mortality outcomes differed by patient age, gender and income deprivation, by estimating fully-interacted models in terms of these patient characteristics (see Appendix Figures 5.B.6, 5.B.7, 5.B.8). The workforce composition shock affected both younger and older patients, although the effect for elderly patients was about three times larger. Female and male patients are both negatively affected, and in a similar scale, by the workforce shock. We find no remarkable differences in the mortality effects for least and most deprived patient.
Furthermore, in order to examine the heterogeneity of the effects of interest by diagnosis type, we also estimated 19 separate static DiD versions of Equation 5.3, one for each of the 19 large patients' subgroups identified by the patient main diagnosis for hospital admission.[25] The results, reported in Appendix Figures 5.B.9, and 5.B.10 and Appendix Table 5.B.9), show that the mortality effects appear to be driven by an increase in deaths for respiratory systems diseases, as well as neoplasms and blood diseases. As respiratory infections require care-intensive treatments by nurses rather than doctors (Fetzer et al., 2024, Friedrich and Hackmann, 2021), it is not surprising that shocks to the nurse workforce composition may be particularly harmful for these patients.

Heterogeneity by Share of 'Leave' Votes in Brexit Referendum. Exploiting the large variation in the share of 'Leave' votes across English counties, ranging from around 20% to 70%, we also investigate the heterogeneity of the in-hospital mortality effects for hospital organizations located in areas with a lower or higher share of 'Leave' votes in the Brexit referendum. We split NHS hospital organizations in 'Remain' and 'Leave' areas (identified by a share of 'Leave' votes respectively below or above 50%) based on the headquarter postcodes of NHS hospital organizations. Given the rather evenly split result of the referendum (52% votes in favour of 'Leave'), there were respectively 47 and 84 NHS hospital organizations in the 'Remain' and 'Leave' areas. Figure 5.B.11 shows that the effects on in-hospital mortality were quite similar for NHS hospital organizations in 'Remain' and 'Leave' areas, although the mortality effects in 'Leave' areas are more precise, probably due to the stronger decline in the share of EU nurses in regions with a high 'Leave' share (see Figure 5.B.12).

5.5.4 Robustness Checks

Exclusion of London Hospitals. London is an outlier in the exposure measure, due to a high density of EU nurses before the referendum, thus the inclusion of London hospitals in the sample may pose a risk to the generalizabity of our main findings. Moreover, the London hospital patient composition might have changed after the referendum vote, given the large number of EU nationals in London. Therefore, as a first robustness analysis we re-estimate our event-study model on a sample of only non-London NHS hospital organizations. When

[24]Respectively: £18,431,700 divided by either £21,909 (Pay Band 5), or £26,302 (Pay Band 6) or £31,383 (Pay Band 7), according to `https://www.rcn.org.uk/employment-and-pay/NHS-pay-scales-2016-17`.

[25]The 19 subgroups are created based on the International Classification of Diseases 10$^{\text{th}}$ Revision (ICD-10) chapter of the main diagnosis code in HES APC.

London hospitals and their patients are excluded from the sample (Appendix Table 5.B.10), the effects on mortality (in-hospital or anywhere) are similarly precise but even larger in magnitude, showing that our main findings are not driven by a London-driven hospital workforce composition effect.

Exposure Variable Measured in the Pre-referendum Year Only. We re-estimate our model on hospital care quality with an exposure variable defined according to the nurse nationality distribution in the last pre-referendum year only, rather than the entire pre-referendum time period. Our results remain substantially unchanged (see Appendix Figure 5.B.13 and Appendix Table 5.B.11).

Controlling for the Exposure to Pre-referendum Share of Non-EU Foreign Nurses. We also estimate an augmented version of Equation 5.3 which includes event-study terms for the pre-referendum exposure to non-EU nurses (Appendix Table 5.B.12). This approach is indicative of whether the effect on quality was driven by the net loss of EU nurses or also by the pre-referendum exposure to a different group of non-native nurses (i.e. non-EU nurses). We show that, to the largest extent, the mortality and readmission effects exclusively depend on the pre-referendum exposure to EU nurses, so that the net loss of EU nurses is the driving force. Since the reduction in EU nurses was primarily driven by the reduction in new joiners from South European countries, we estimate four different models each having a pre-referendum exposure measure computed among nurses coming from a specific EU national subgroup. Consistently with the findings on the nurse composition changes, Appendix Figure 5.B.14 shows that in-hospital mortality was more affected by the exposure to South European, East European and Irish nurses, and never by the pre-referendum share of North European nurses, which was largely unaffected by the Brexit vote shock.

Controlling for the Exposure to Pre-referendum Share of EU Doctors. Furthermore, we estimate an event-study specification in which we add the pre-referendum exposure to EU senior hospital doctors – computed in the same fashion as for nurses. As there was no drop in EU doctors after the referendum, we would expect mortality effects to be caused mostly by the nurse workforce composition shock. Therefore, these results provide us also with a first useful falsification test. The estimates, displayed in Appendix Table 5.B.13 and Figure 5.B.15, show that the effects on patient mortality and readmission rates are clearly associated with the pre-referendum exposure to EU nurses and not EU doctors.

Falsification Tests Using Randomization Inference. Finally, we also perform a falsification exercise based on randomization inference with 300 replications. For the continuous placebo exposure variable, we simulated 300 random draws from a log-normal distribution with the same mean and variance of the original pre-referendum hospital share of EU nurses; instead, for the binary placebo exposure variable, we 300 random values from a uniform distribution and created a binary exposure indicator equal to one for uniform draws over 0.75, and zero otherwise. We assigned the aforementioned placebo exposure variables to the 131 NHS hospital organizations in our sample, and then estimate the event-study regressions as in Eq. 5.3 on the observed patient health outcomes and control covariates, based either on the continuous or the binary placebo exposure variables.

The complete distributions of the 13 event study pre- and post-referendum coefficients are reported in Appendix Figures 5.B.17 and 5.B.18: the distribution of each estimate is symmetric and centered around zero. Appendix Figure 5.B.16, instead, shows the event-study plot obtained by reporting the average point estimates, upper and lower bounds of the 95% confidence intervals, from the distribution the 300 event-study regressions. Also in this case, the estimates of the placebo exposures are centered around zero, both before and after the referendum, providing additional evidence that our main findings are not due to chance.

5.6 Mechanisms

In this section, we provide empirical evidence on the possible mechanisms explaining the deterioration in care quality in hospitals exposed to the Brexit referendum shock. In the first instance, we test the mechanism proposed by our conceptual framework, that is, the hospital nurse workforce composition changes triggered by the Brexit referendum resulted in a decrease in the skill level of newly hired NHS hospital nurses. Consistently with our theoretical predictions, we show that the quality of the newly joiner nurses after the Brexit referendum is worse than that of the nurses who joined the NHS before the referendum. Subsequently, we test a series of alternative or complementary mechanisms to the skill deterioration that we proposed, and we find no evidence that they explain the quality effects of interest.

5.6.1 Changes in Nurses' Skills Composition

We use the first observed pay grade that foreign nurses are assigned to, when they join their NHS hospital employer, as a measure to gauge the skill levels of newly hired foreign nurses by NHS hospitals, before and after the referendum.[26] This strategy is based on the following facts and assumptions: i) for most job hiring decision, pay is a good indicator of workers' skills and human capital at the aggregate level, i.e. on average for a group of workers; ii) a global labour market for hospital nursing jobs existed already in the years before the Brexit referendum; iii) given the international mobility of foreign nurses, the joiner nurses, who were hired after the Brexit referendum, would have chosen to move to a different employer, or to a country different from England, if they considered the pay package offered by their

[26] Using the observed pay grade variable reported in ESR records has several advantages: it is an objective measure for the expertise of nurses, and so a plausible proxy for their skill level; it is observed (i.e., non-missing) for all newly hired nurses; last but not least, it is not model-based, thus not prone to introduce model and/or measurement errors in this analysis. Instead, we cannot employ a Abowd-Kramartz-Margolis approach (Abowd et al., 1999) (hencefort, AKM) to measure newly hired nurses' skills, for several reasons. First, we are interested in measuring nurses' skills right at the moment of their NHS hospital entry, and not after they have acquired additional human capital through specific or general training at the NHS hospital where they are hired, but we do not have any record of the tasks, activities and qualifications that the joiner nurses possess before they are firstly employed by the a NHS hospital. Second, in HES APC there is no records of the nursing team members in charge of a given patient. Third, a hypothetical AKM strategy would require estimating separate matched nurse-hospital fixed effects regression models in the years before and after the Brexit referendum, but the estimated fixed effects risk to be biased due to the structural break induced by the Brexit referendum shock.

NHS hospital employer too low for their skill level; iv) the existence of a different NHS 'pay policy' for foreign nurses, i.e. hiring nurses in higher (lower) pay grades before (after) the referendum, is extremely unlikely.[27]

Therefore, we analyze how the pay bands of foreign newly joiner nurses changed, before and after June 2016. Higher pay bands reflect better qualifications, better employment references and a higher pre-employment tenure, that is, a greater job experience. Figure 5.7 shows the share of nurses being employed in wage bands 1-4 (*panel a.*), wage band 5 (*panel b.*), and wage band 6 (*panel c.*) among all newly joining, foreign nurses. After the Brexit referendum, the share of foreign joiner nurses employed in wage band 1 to 4 increased and doubled from 2016 to 2019. The share of foreign joiner nurses in the higher wage bands, proxying higher quality, decreased. ESR records also provide us with NHS hospital employees' salary spinal point, which gives the most precise measure of the nurse basic salary and pay level in a month or year; also when we use this more accurate observed pay level, with respect to pay bands, we find that the average salary spinal point (*panel d.*) of foreign newly joiner nurses fell after the referendum . Similarly, the minimum (*panel e.*) and maximum salary (*panel f.*) within the nurses' grades fell, mirroring the findings on the wage band structure.[28]

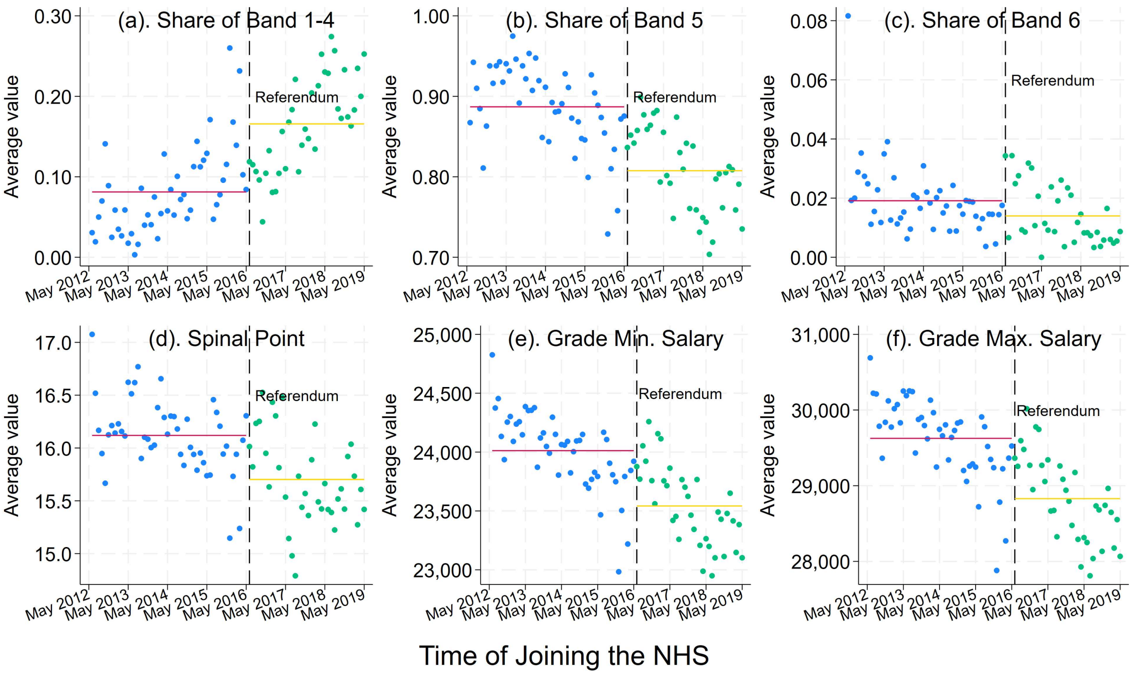
Hence, after Brexit, the composition of foreign nurses deteriorated given the increasing inflow of low-wage band nurses. This is a consequence of hiring more non-EU joiners who on average are hired at lower wage bands and lower spinal points (see Appendix Figures 5.B.19 and 5.B.20, and the related pre- and post-referendum averages reported in Table 5.B.15). This matches the propositions derived in our theoretical model.

Additionally, we exploit yearly National Staff Survey data, for the years 2012 to 2019, to investigate if nurses' self-reported satisfaction with the quality of care they provide to patients and their working environment changed, based on the share of pre-referendum EU nurses. The nurse satisfaction outcome variables in the surveys are expressed on a Likert scale from 1 to 5 (e.g. strongly disagree; disagree; neither agree nor disagree; agree; strongly agree). For this reason we estimate ordered logit regressions, including hospital organization fixed effects and year-NHS region fixed effects, to mimic the event-study specification Equation 5.3 used to evaluate the effect on patient health outcomes. We provide event-study models for two samples of nurses. The first sample includes all the nurses employed in the NHS hospitals and responding to the survey; the second sample, instead, consists only of British nurses who have been employed for at least six years in a hospital, which ensures that our results are not affected by selection (e.g. joining and/or leaving EU nurses might differ from joining

[27]The last assumption is justifiable according to two simple considerations. First, the NHS is highly unionized and also equal opportunity employer, so such a 'discriminatory' policy would have generated highly heated political and media debates, which we are not aware of. Secondly, such a policy would be inconsistent with labour market forces and the chain of events that we have shown: after the referendum the NHS had to cope with urgency the recruitment of foreign nurses, to avoid hospital nurse shortages due to the missing inflow of EU joiner nurses, so we should expect that highly-demanded foreign nurses should have experienced a relative increase in their wage-bargaining power, with respect to the pay grade that they were assigned to when joining NHS hospitals.

[28]The average values displayed in Panels e and f of Figure 5.7 measure the minimum and maximum salary levels that can be potentially earned by newly joining nurses, given their starting employment grade. Thus, the decreases highlighted by the two figures further confirms a reduction in the average employment grade among new hires.

Figure 5.7: Changes in foreign joining nurse composition by pay band and salary grade



Note: This figure gives the composition of joiners by wage group, spinal point, or minimum and maximum salary over time.

non-EU nurses) or the fact that nurses were directly affected by the referendum (e.g. due to worsening residence regulations).

Table 5.2 reports coefficients in odds ratios, and it shows that the likelihood of nurses reporting that they looked forward to going to work significantly decreased (odds ratios lower than one) in hospitals more exposed to Brexit, compared to less exposed hospitals. This effect is especially prominent among British nurses who have been employed before and after the referendum. Moreover, after the referendum the nurse satisfaction with their own quality of care, as well as the ability to deliver care as aspired, decreased more in hospitals with higher exposure to the Brexit shock. Finally, fewer British nurses felt like they made a difference to their patients. We consider the findings above as evidence that nurses employed by NHS hospitals heavily exposed to the Brexit shock perceived a deterioration in the quality of patient care they provided.

5.6.2 Alternative and Complementary Mechanisms

There could be alternative or complementary mechanisms at play that may explain the deterioration of hospital care quality that we document. Hereafter we provide evidence on four channels, either on the hospital care supply or demand sides.

Nursing Workforce Shortages. A first check on the hospital supply side is whether the exposure to the referendum shock had any effect on nurse labour capacity, that is, generated nurse shortages. According to Table 5.1, there was no effect on the number of nurses employed in the hospital – hence, no shortages of nurses in hospitals with a higher pre-2016

Table 5.2: Effects on nurse self-reported satisfaction with provision of hospital care

	I look forward to going to work.	I am satisfied with the quality of care I give to patients / service users.	I am able to deliver the care I aspire to.	I feel that my role makes a difference to patients / service users.
	(1)	(2)	(3)	(4)
Panel A: All nurses				
I(2013 NSS) * Pre-BR share of EU nurses	0.998			
	[0.975,1.021]			
I(2014 NSS) * Pre-BR share of EU nurses	1.011			
	[0.995,1.026]			
I(2016 NSS) * Pre-BR share of EU nurses	1.000	0.999	0.998	1.001
	[0.988,1.013]	[0.983,1.015]	[0.978,1.017]	[0.989,1.013]
I(2017 NSS) * Pre-BR share of EU nurses	0.990	0.977**	0.983*	0.989*
	[0.977,1.003]	[0.958,0.996]	[0.963,1.002]	[0.977,1.002]
I(2018 NSS) * Pre-BR share of EU nurses	0.994	0.987	0.988	0.990
	[0.979,1.009]	[0.967,1.007]	[0.967,1.009]	[0.977,1.003]
I(2019 NSS) * Pre-BR share of EU nurses	0.987	0.980**	0.982*	0.990
	[0.971,1.004]	[0.962,0.999]	[0.963,1.000]	[0.976,1.004]
Observations (nurse responses to NHS Staff Surveys)	398,953	333,969	333,161	333,435
Panel B: British nurses employed by at least 6 years				
I(2013 NSS) * Pre-BR share of EU nurses	0.991			
	[0.971,1.010]			
I(2014 NSS) * Pre-BR share of EU nurses	1.011			
	[0.992,1.030]			
I(2016 NSS) * Pre-BR share of EU nurses	0.994	0.997	1.000	0.998
	[0.978,1.011]	[0.978,1.016]	[0.977,1.023]	[0.983,1.013]
I(2017 NSS) * Pre-BR share of EU nurses	0.982**	0.968***	0.974**	0.982**
	[0.967,0.998]	[0.946,0.990]	[0.952,0.997]	[0.968,0.997]
I(2018 NSS) * Pre-BR share of EU nurses	0.985*	0.981	0.982	0.990
	[0.968,1.003]	[0.960,1.004]	[0.958,1.006]	[0.975,1.004]
I(2019 NSS) * Pre-BR share of EU nurses	0.977**	0.972***	0.978**	0.981***
	[0.959,0.995]	[0.952,0.992]	[0.959,0.997]	[0.968,0.994]
Observations (nurse responses to NHS Staff Surveys)	196,453	161,208	160,886	160,990

Note: $N_{clusters}$ = 131 acute care NHS hospital organizations. Period: 2013-2019. Outcome variables: nurse responses to yearly NHS Staff Surveys, expressed on a 1–5 Likert scale. Ordinal logit odds ratios and corresponding 95% confidence intervals. Robust standard errors, clustered at the hospital organisation level. The event-study specification is based on Equation 5.3, using a continuous exposure to the Brexit shock. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

share of EU nurses.

Financial Capacity. Another concern is whether more exposed NHS hospital organizations suffered a change in their financial conditions after the referendum. If hospital organizations more exposed to the Brexit shock also experienced a surge in patient costs or a fall in their revenues, they might have skimped on patient care and safety measures to revert their dire financial situation. To investigate this mechanism, we use the publicly available financial accounts data at NHS hospital organization level (referenced in Section 5.2.3) and estimate event-study specifications in which we use the natural logarithm of total hospital expenditures and revenues as dependent variables.[29] The estimates, provided in Appendix Figure 5.B.21, show that there was no Brexit effect either on the income or the expenditures of the hospitals more exposed to the migration shock induced by the referendum.

Bed Occupancy. While the overall number of nurses was unaffected, and we find no apparent evidence of nurse shortages in hospitals more exposed to the Brexit shock, it is possible that labour productivity in such hospitals decreased compared to the pre-referendum period.

[29]For several NHS hospital organizations, the published financial accounts data is unavailable in some years of the sample, and therefore missing; as such, the results of this analysis is based on an unbalanced sample consisting in 85 out 131 of the hospital organizations in the main sample.

To test this hypothesis, we analyze how the exposure to the Brexit shock relates to hospital bed occupancy rates.[30] Appendix Figure 5.B.22 shows that the bed occupancy rate of hospitals with a higher pre-referendum share of EU nurses suffered a modest decrease of about 0.50% in the post-referendum period, with respect to the 86.88% pre-referendum mean bed occupancy rate, but this decrease is not significant even at 10% level. Provided that the bed occupancy rate is a reliable measure of hospital productivity, these results point towards a possible weak decrease in the productivity of hospital organizations more exposed to the Brexit shock. This finding is likely consistent with changes in the hospital (nursing) workforce composition, as highlighted in Section 5.6.1, that might have averted labour shortages through the imperfect substitution of the missing inflow of EU nurses.[31]

Changes in Hospital Patient Demand. Finally, we investigate whether our main effects on quality of care might be explained by changes in hospital patient demand across different English NHS hospitals. We investigate whether, across hospitals differently exposed to the shock, there were changes in the population of the NHS provider catchment area.[32] The catchment area is calculated by Public Health England based on the frequency of patient hospital utilization and admissions. If patients switched hospitals of treatment or strategically changes their residence, e.g. anticipating a lower quality of care in the hospitals more exposed to the foreign labour shock, the catchment area population would decrease. Appendix Figures 5.B.23 and 5.B.24 show that there was no significant change in the catchment area population for all age groups and at the aggregate level. Moreover, Appendix Table 5.B.14 shows that the total number of patients admitted to the hospital was unaffected by the exposure variable.[33] Hence, reverse patient mobility, choice and utilization do not appear as plausible mechanisms that can explain the findings on quality of hospital care.

5.7 Conclusions

In many developed countries like the UK, several sectors of the economy are critically reliant on the immigration of skilled workers. Our work provides several insights to the existing economics literature on this matter, drawing from the relevant case of migrant nurses who are employed in the English NHS hospital acute care sector.

Skilled migrant workers like nurses are responsive to changes in the institutional settings and hospitality environment of prospective hosting countries: we document how the outcome of the 2016 Brexit referendum led to a significant change in the nursing workforce composition

[30] For this analysis, we use NHS bed occupancy data publicly available at: `https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-data-overnight/`.

[31] If hospital bed occupancy rates increased in the more exposed hospital organizations after the referendum, consistently with a quality-volume trade-off mechanism, this channel would represent an alternative explanation to our preferred mechanism. However, as the sign of the (non significant) effect on the hospital bed occupancy rate is negative, our event study estimates find no empirical support in favour of this alternative mechanism.

[32] Data on the catchment area of NHS providers are publicly available at: `https://app.powerbi.com/view?r=eyJrIjoiODZmNGQ0YzItZDAwZi00MzFiLWE4NzAtMzVmNTUwMThmMTVlIiwidCI6ImVlNGUxNDk5LTRhMzUtNGIyZS1hZDQ3LTVmM2NmOWRlODY2NiIsImMiOjh9`.

[33] For this analysis we use the aggregate data on the volume of patients admitted to NHS hospital organization during the years of our sample period, recorded in HES APC data.

in NHS acute care hospitals, driven by a stark decrease in the number of EU joiner nurses. This evidence is in line with economic theory: prospective foreign workers decide whether to move to work in a given host country, form expectations based on the existing labour market conditions, and revise such expectations when big shocks to the labour market arise, as in the case of the Brexit referendum. Instead, migrant workers already employed in the host country may be much less responsive to migration policy shocks, such as the Brexit referendum outcome, for several reasons: they may have already gained settlements rights; they may postpone their reactions to the moment when the new immigration regulatory framework is clearly defined along with their settlement rights (in our case study, this would have been the approval of the 2020 European Union Withdrawal Agreement Act by the UK Government); or their reactions are smoothed over a longer time window due to the expensiveness of an otherwise abruptly quick divestment process to leave the UK.

Moreover, we find that sudden changes in the composition of skilled workers have the potential to disrupt the quality of healthcare services provided. In the case we studied, patients admitted to a NHS hospital organization with an average exposure to the Brexit referendum shock experienced a 5.31% increase in the risk of in-hospital death, and 2.28% increase in the risk of unplanned emergency hospital readmission, after June 2016. This translates into about 1,485 additional in-hospital deaths per year, in the three years after the referendum, or equivalently 2.18 extra in in-hospital deaths for each of the fewer 2,047 EU nurses employed in English NHS hospital organizations after the referendum. The size of these Brexit-related mortality effects are therefore quite large, especially when we compare them with the mortality impact of a catastrophic event such as the COVID-19 outbreak (about 3,050 non-COVID-related deaths, according to Fetzer et al. 2024): as the Brexit referendum was a scheduled political event, whereas the COVID-19 pandemic outbreak was likely unavoidable, English healthcare policy-makers could have put in place contingency plans to attenuate the Brexit shocks due to changes in the NHS hospital workforce composition, in case of a 'Leave' victory scenario at the referendum.

We also find that the risk of unplanned emergency readmission to hospital increased by about 2.78% in hospital organizations more exposed to the Brexit shock, equivalently to about 8,777 additional unplanned readmissions. Despite an unplanned emergency readmission is a much less severe event than the absorbing case of a patient death, it may be still a very stressful event for the patients' physical and mental health, and it produces extra work burden for overworked and fatigued NHS nurses and doctors; it also results in £18.4 millions extra costs for the NHS, which could have been employed to hire for one year 841 nurses, approximately equivalent to half of the nursing staff of an average size NHS hospital organization.

These empirical results can be reconciled through the lens of the theoretical framework that we provide, and the workforce composition mechanism that we test empirically: the most readily available nurses to start a job in NHS hospitals with short notice would have likely been exactly those nurses with lower reservation wages or opportunity-costs from leaving another nursing job elsewhere, in the UK or abroad. As such, the workforce composition changes in the NHS hospital nursing workforce may have prevented the insurgence of long-lasting and severe nurse shortages, but at the cost of a decrease in the quality of new hires: we find suggestive evidence supporting this mechanism by analyzing the changes in both the pay grades of new joiner nurse hired by NHS hospital, and also in the level of nurses' satisfaction with the quality of services they provide.

Overall, our investigation emphasizes the importance of high-skilled, foreign nurses in hospital care, and contribute more generally to understand the effects of workforce composition and of foreign labour supply extensive margins on the performance of labour-intensive organizations such as public hospitals. The takeaway message from this study is that, in countries relying on skilled foreign labour force, such as the US and the UK, political initiatives fostering nationalistic interests and with a relevant expected impact on immigration patterns should be carefully weighed against the potential disruptions to the labour supply chain in critical sectors of the economy, such as health care. These detrimental effects should not be underestimated in labour-intensive sectors, such as health care, also in light of the ongoing demographic changes and the ever-increasing demand for skilled (healthcare) workers in highly developed countries.

Ultimately, our research suggests that policy-makers should take informed decisions based on the willingness to move of prospective native and foreign skilled workers in the short, medium and long-term, according to different immigration scenarios. Failing to do so can critically disrupt organizational performance, at the very least in the sectors mostly exposed to immigration-related labour supply shocks.

Appendix A: Proofs

Proof of Lemma 1. Let $h^*(\theta)$ be the hiring decision described in the Lemma's statement. That is, $h^*(\theta) = 1$ if $\theta \geq \theta^*$ and zero otherwise. Let $h : \mathbb{R} \to [0, 1]$ be an arbitrary hiring rule satisfying condition (5.2). We divide the proof into two cases: $\tilde{\theta} < \theta_0$ or $\tilde{\theta} \geq \theta_0$.

Case I: Suppose $\theta^* = \theta_0 > \tilde{\theta}$. The difference in total quality under h^* compared to h is

$$Q(h^*, a) - Q(h, a) = -\int_{-\infty}^{\theta_0} q(\theta)h(\theta)a(\theta)d\theta + \int_{\theta_0}^{+\infty} q(\theta)[1 - h(\theta)]a(\theta)d\theta \geq 0.$$

The first term on the right-hand-side is positive because $q(\theta) \leq 0$ for all $\theta \leq \theta_0$. The second term is positive since $q(\theta)[1 - h(\theta)]a(\theta) \geq 0$ for all $\theta > \theta_0$.

Case II: Suppose $\theta^* = \tilde{\theta} \geq \theta_0$. The difference in total quality under h^* compared to h is

$$Q(h^*, a) - Q(h, a) = -\int_{-\infty}^{\theta_0} q(\theta)h(\theta)a(\theta)d\theta + \int_{\theta_0}^{+\infty} q(\theta)[h^*(\theta) - h(\theta)]a(\theta)d\theta \geq 0.$$

The first term on the right-hand-side is positive because $q(\theta) \leq 0$ for all $\theta \leq \theta_0$. For the second term, note that $[h^*(\theta) - h(\theta)]a(\theta)$ is never strictly positive then strictly negative. Moreover, q is increasing and $q(\theta) \geq 0$ for all $\theta > \theta_0$ and $\int_{\theta_0}^{+\infty} h^*(\theta)a(\theta)d\theta = M \geq \int_{\theta_0}^{+\infty} h(\theta)a(\theta)d\theta$. Hence, the second term is also positive by the Beesack's inequality (Beesack (1957)). □

Proof of Proposition 1. We prove each of the items separately.

Item 1: Suppose for the sake of obtaining a contradiction that $\theta^*_{pre} < \theta^*_{post}$. If that is the case, then (5.2) must also bind after the referendum, which implies that the total mass of hired workers before and after the referendum must be the same. That is,

$$\int_{\theta^*_{post}}^{+\infty} a_{post}(\theta)d\theta = \int_{\theta^*_{pre}}^{+\infty} a_{pre}(\theta)d\theta,$$

where a_{pre} and a_{post} denote the mass of applicants of each type before and after the referendum. Note, however, that as $\gamma_e^{pre} > \gamma_e^{post}$ we have that

$$\int_{\theta^*_{post}}^{+\infty} a_{post}(\theta)d\theta - \int_{\theta^*_{pre}}^{+\infty} a_{pre}(\theta)d\theta =$$
$$\int_{\theta^*_{post}}^{+\infty} \mu \underbrace{\left[F(\omega_e + \gamma_e^{post}) - F(\omega_e + \gamma_e^{pre})\right]}_{<0} g(\theta)d\theta - \int_{\theta^*_{pre}}^{\theta^*_{post}} \underbrace{a_{pre}(\theta)}_{>0} d\theta < 0.$$

A contradiction. □

Item 2: The total number of workers hired cannot increase post-referendum since (5.2) was binding pre-referendum. However, the number of non-EU hired workers increases since $\theta^*_{post} < \theta^*_{pre}$. Therefore, the share of newly hired EU workers decreases. □

Item 3: Note that

$$Q(h^*_{pre}, a_{pre}) - Q(h^*_{post}, a_{post}) = \int_{\theta_0}^{+\infty} q(\theta)\left[h^*_{pre}(\theta)a_{pre}(\theta) - h^*_{post}(\theta)a_{post}(\theta)\right]d\theta.$$

Recall that q is increasing and $q(\theta) > 0$ for all $\theta > \theta_0$.

Moreover, $\left[h^*_{pre}(\theta)a_{pre}(\theta) - h^*_{post}(\theta)a_{post}(\theta)\right]$ single-crosses zero from below and

$$\int \left[h^*_{pre}(\theta)a_{pre}(\theta) - h^*_{post}(\theta)a_{post}(\theta)\right]d\theta \geq 0.$$

Therefore, by the Beesack's inequality $Q(h^*_{pre}, a_{pre}) > Q(h^*_{post}, a_{post})$. $\square$

Item 4: We define worker shortages as not all vacancies being filled, or equivalently, $\theta^* = \theta_0 > \tilde{\theta}$. Note that, by Lemma 1, θ^* is a continuous and, by item 1, decreasing function of γ_e. Moreover, if $(\gamma_e^{pre} - \gamma_e^{post}) = 0$, then $\theta^*_{post} = \theta^*_{pre} > \theta_0$. Therefore, $\theta^*_{post} > \theta_0$, unless the decrease in $(\gamma_e^{pre} - \gamma_e^{post})$ is sufficiently large.$\square$ $\square$

Proof of Proposition 2. The mass of unfilled vacancies (if any) plus the number of workers hired after the referendum with skills below the pre-referendum cutoff is equal to the mass of prospective workers with type above θ^*_{pre} who would apply pre-referendum and no longer do (for instance, areas A and D in figures 5.1a and 5.1b). That is,

$$\begin{aligned}\int_{\tilde{\theta}_{post}}^{\theta^*_{pre}} a_{post}(\theta)d\theta &= \underbrace{\int_{\tilde{\theta}_{post}}^{\theta^*_{post}} a_{post}(\theta)d\theta}_{\text{Unfilled vacancies}} + \underbrace{\int_{\theta^*_{post}}^{\theta^*_{pre}} a_{post}(\theta)d\theta}_{\text{Hired below } \theta^*_{pre}} \\ &= \mu \int_{\theta^*_{pre}}^{+\infty} \Big[F(\omega_e + \gamma_e^{pre}) - F(\omega_e + \gamma_e^{post})\Big] g(\theta)d\theta.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{d\int_{\tilde{\theta}_{post}}^{\theta^*_{pre}} a_{post}(\theta)d\theta}{d\mu} &= \int_{\theta^*_{pre}}^{+\infty} \Big[F(\omega_e + \gamma_e^{pre}) - F(\omega_e + \gamma_e^{post})\Big] g(\theta)d\theta \\ &\quad - \mu\Big[F(\omega_e + \gamma_e^{pre}) - F(\omega_e + \gamma_e^{post})\Big] g(\theta^*_{pre})\frac{d\theta^*_{pre}}{d\mu}.\end{aligned}$$

As $\gamma_e^{pre} > \gamma_e^{post}$, the first term of the right-hand-side is positive. Hence, if we show that $d\theta^*_{pre}/d\mu < 0$ we are done. Recall that

$$M = \int_{\theta^*_{pre}}^{+\infty} \Big[\mu F(\omega_e + \gamma_e^{pre}) + (1-\mu)F(\omega_r + \gamma_r)\Big] g(\theta)d\theta.$$

Totally differentiating with respect to μ and isolating $d\theta^*_{pre}/d\mu$, we get

$$\frac{d\theta^*_{pre}}{d\mu} = \frac{\int_{\theta^*_{pre}}^{+\infty} \left[F(\omega_e + \gamma_e^{pre}) - F(\omega_r + \gamma_r)\right] g(\theta)d\theta}{\Big[\mu F(\omega_e + \gamma_e^{pre}) + (1-\mu)F(\omega_r + \gamma_r)\Big] g(\theta^*_{pre})}$$

which is smaller than zero, as $\omega_r + \gamma_r > \omega_e + \gamma_e^{pre}$ and F is strictly increasing. $\square$

Proof of Proposition 3. Note that the hiring skill cutoff θ^* is bounded above by $\overline{\theta} := G^{-1}(1 - M)$. $\overline{\theta}$ would be the hiring cutoff if all potential workers were to apply. The fewer the applicants, the smaller the hiring cutoff. Consider now a sequence $(\omega_{e,n}, \omega_{r,n})$ where both

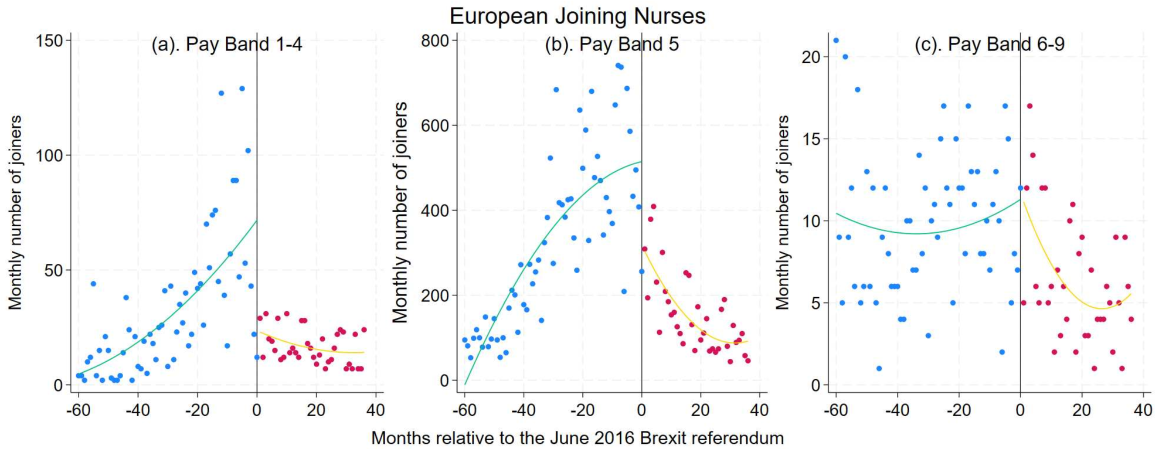
$\omega_{e,n} \to +\infty$ and $\omega_{r,n} \to +\infty$ as $n \to +\infty$. For each $n \in \mathbb{N}$, let $\theta^*_{pre,n}$ and $\theta^*_{post,n}$ be the pre and post-referendum hiring cutoffs associated with a pair $(\omega_{e,n}, \omega_{r,n})$. Hence, for n sufficiently large and $\ell \in \{pre, post\}$ we have

$$M = \int_{\theta^*_{\ell,n}}^{+\infty} \Big[\mu F(\omega_{e,n} + \gamma^{\ell}_{e,n}) + (1-\mu) F(\omega_{r,n} + \gamma^{\ell}_{r,n})\Big] g(\theta) d\theta.$$

As n increases, both $F(\omega_{e,n} + \gamma^{\ell}_{e,n})$ and $F(\omega_{r,n} + \gamma^{\ell}_{r,n})$ converge to one. Hence, both $\theta^*_{pre,n}$ and $\theta^*_{post,n}$ converge to $\overline{\theta}$. Therefore, $|\theta^*_{pre,n} - \theta^*_{post,n}| \to 0$ as $n \to +\infty$. $\square$

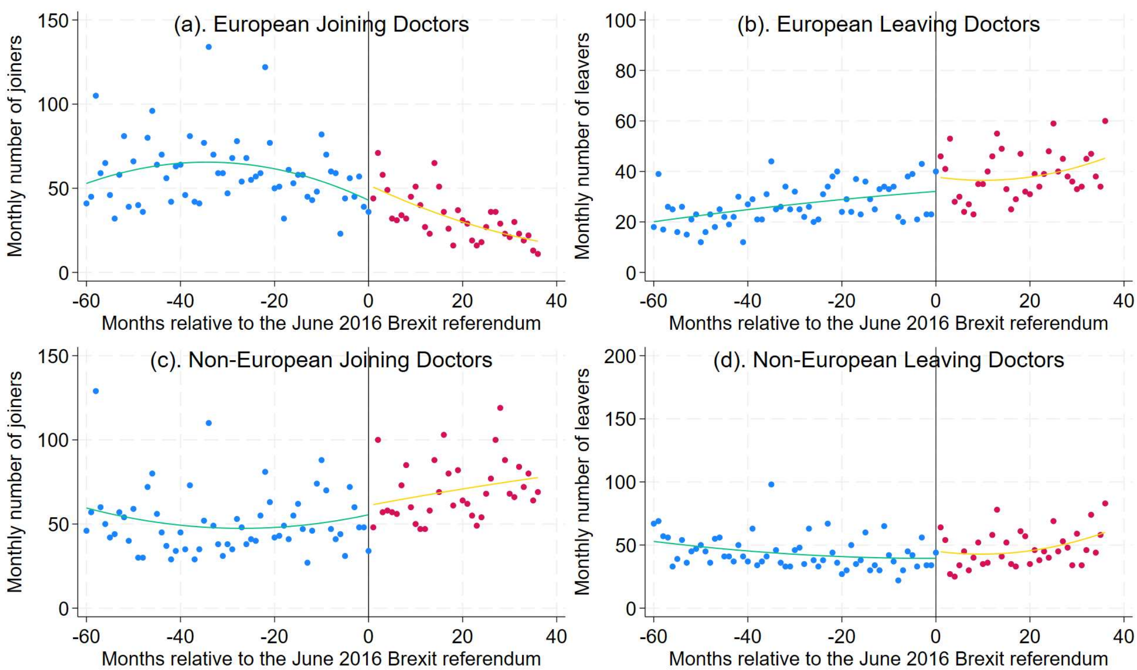
Appendix B: Additional Figures and Tables

Figure 5.B.1: European nurses joining the English NHS by pay band section



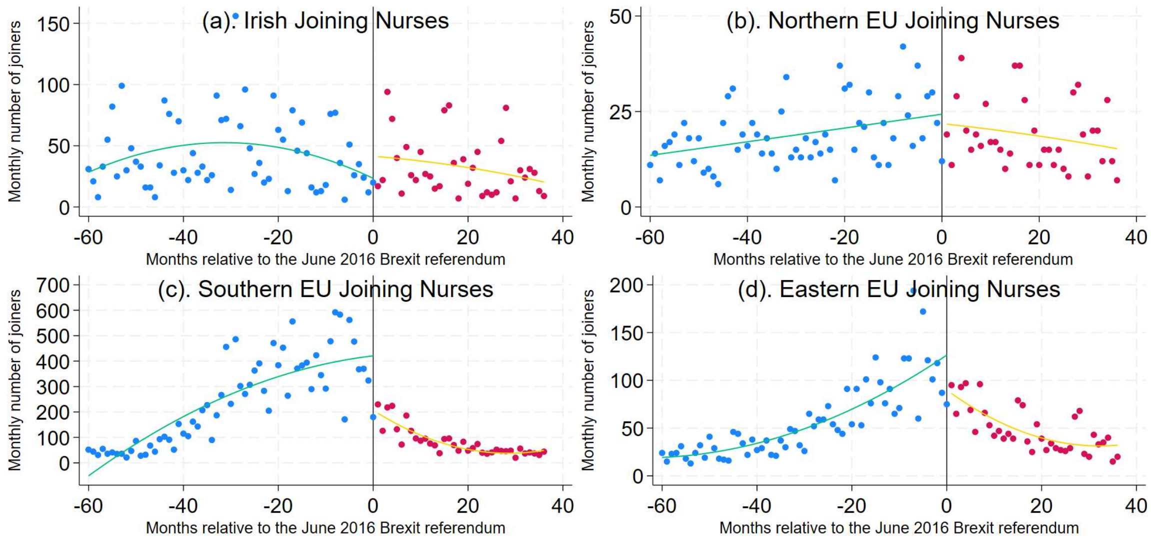
Note: This figure gives the number of monthly European nurses joining by wage band and over time with the Brexit referendum in month 0.

Figure 5.B.2: Senior doctors joining and leaving the English NHS by nationality group



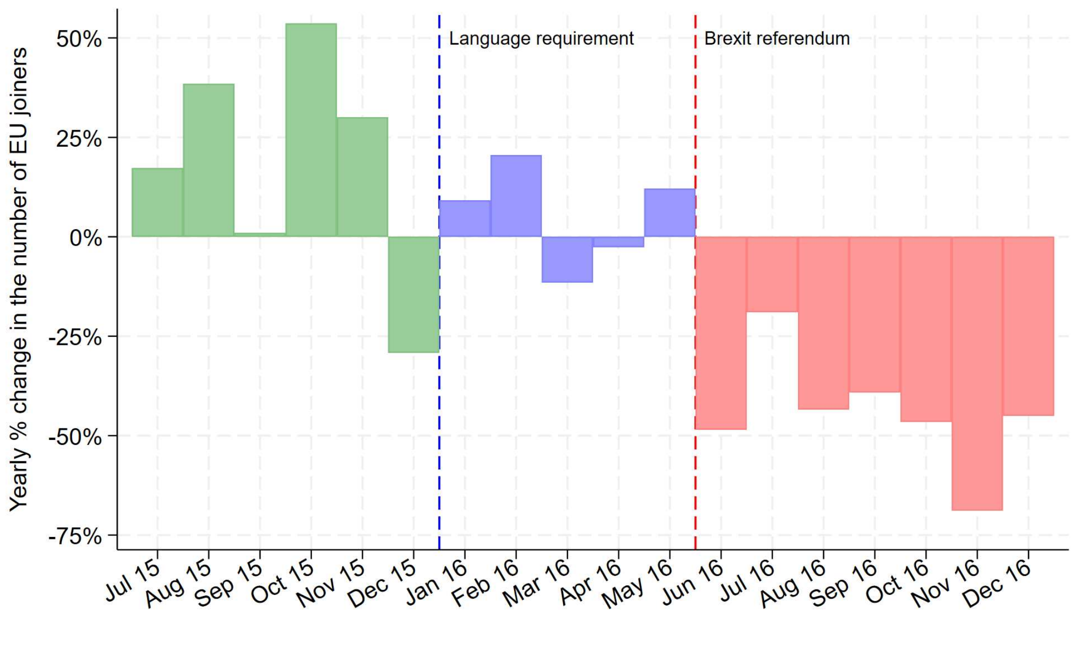
Note: This figure gives the number of monthly European and non-European doctors joining and leaving over time with the Brexit referendum in month 0.

Figure 5.B.3: Nurses joining the English NHS by different European nationality subgroups



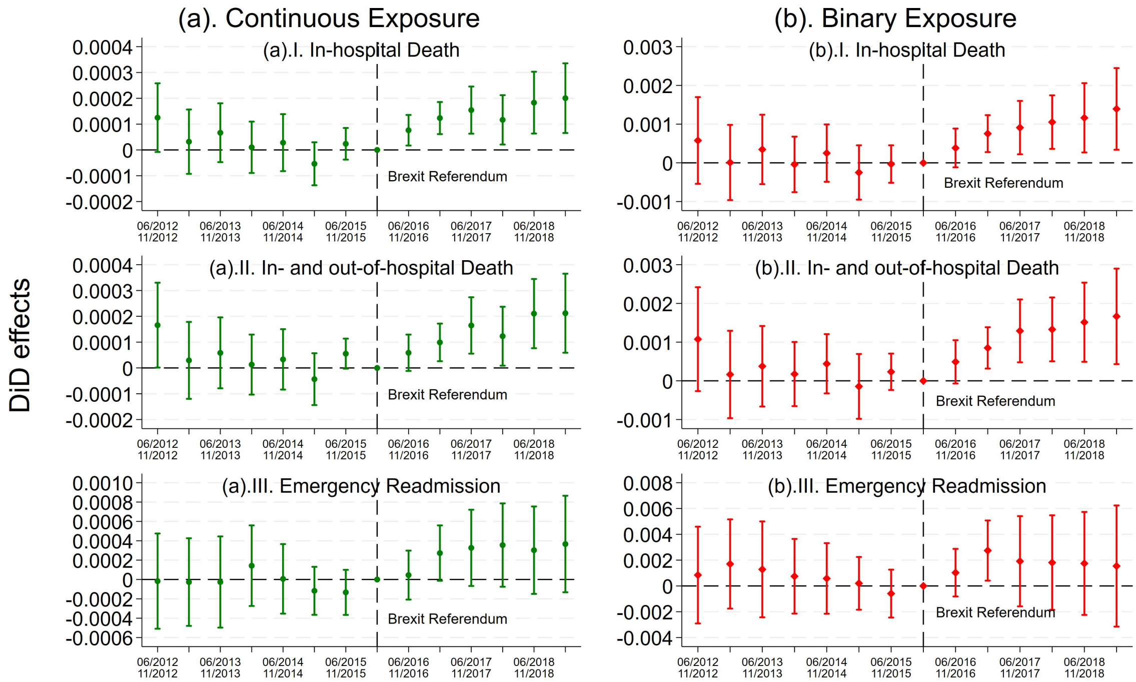
Note: This figure gives the number of monthly joiners to the NHS over time by nationality groups of European nurses with the Brexit referendum in month 0.

Figure 5.B.4: Yearly change in EU nurse joiners around the Brexit referendum date



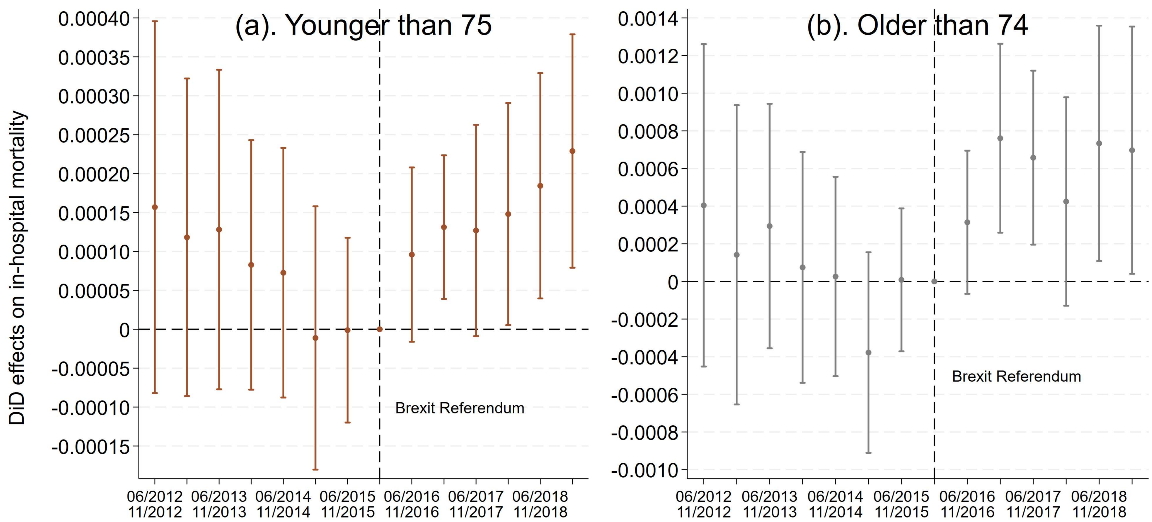
Note: This figures gives the percentage-point change in the total monthly number of NHS nurse joiners from Europe compared to the same month of the year before.

Figure 5.B.5: Dynamic DiD effects of Brexit referendum on Individual Health Outcomes (all patients)



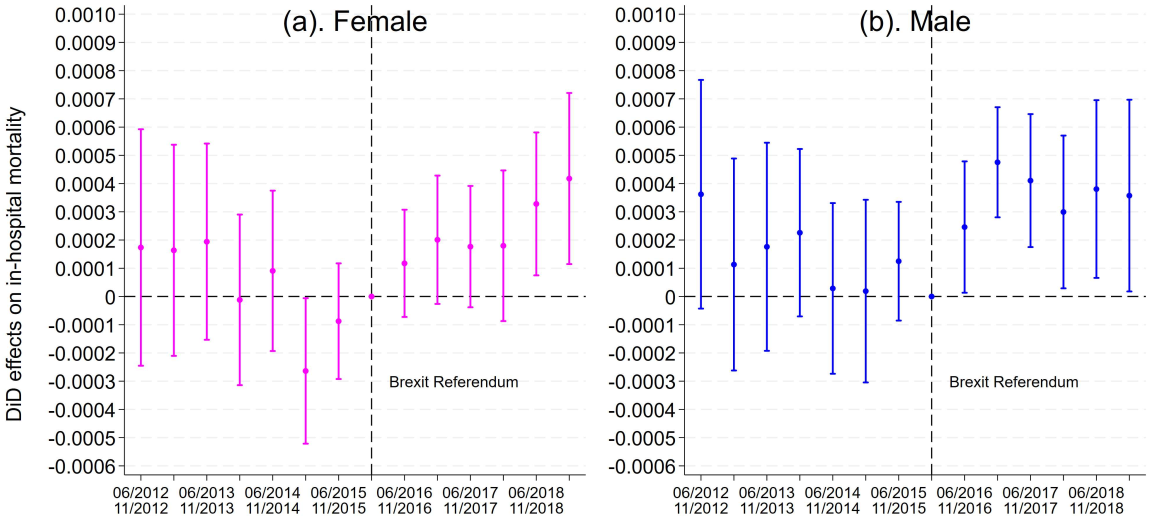
Note: N = 89,728,352 hospital admissions (both emergency and non-emergency). $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating Equation 5.3 with the continuous and the binary (i.e., below/above 75th percentile exposure) exposure. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals. The joint Wald test F-stat values, and their related p-values, testing the null hypothesis that all the leads effects are equal to zero, are reported in Appendix Table 5.B.8.

Figure 5.B.6: Dynamic DiD effects on in-hospital mortality, by emergency patients' age



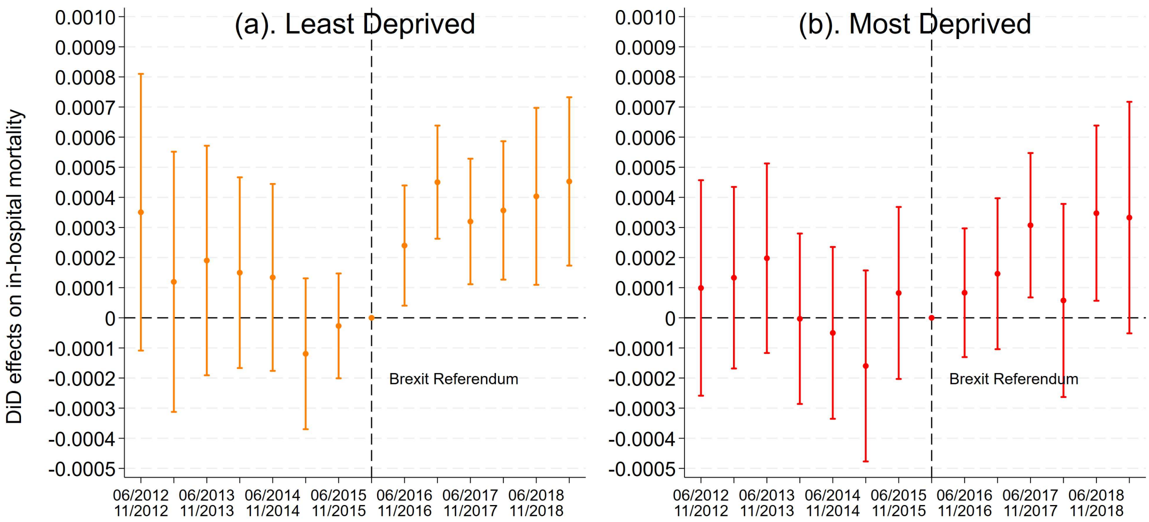
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating the triple-difference version of Equation 5.3 with the continuous exposure interacted with a dummy whether the patient is younger than 75 (age<75) or older than 74 (age≥75). Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.7: Dynamic DiD effects on in-hospital mortality (emergency patients), by gender



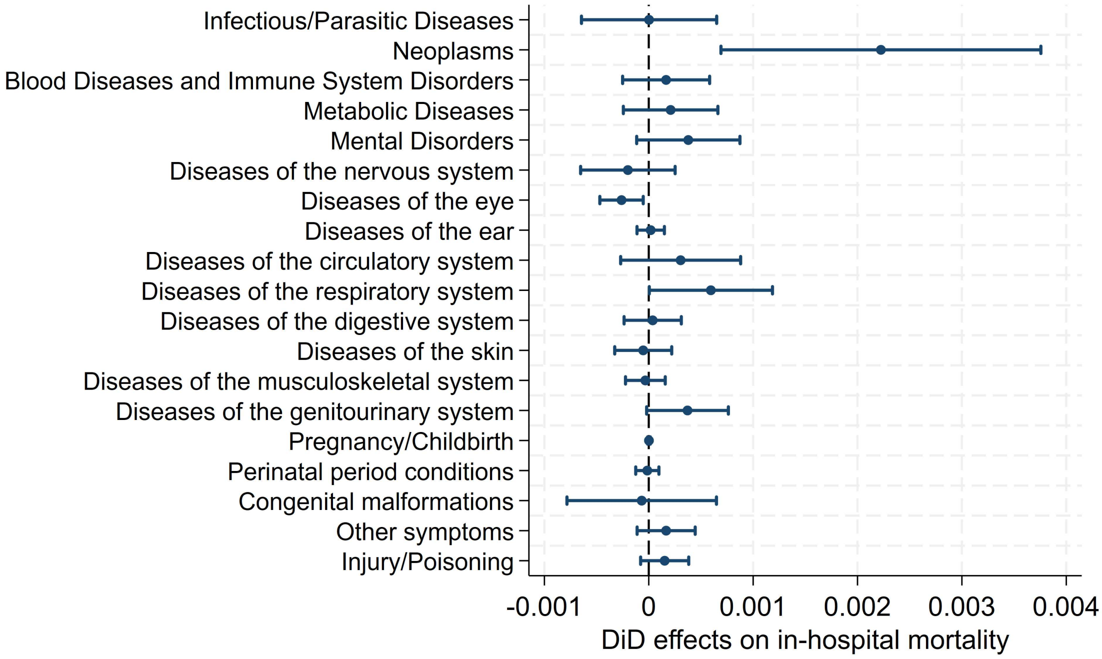
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating the triple-difference version of Equation 5.3 with the continuous exposure interacted with a dummy whether the patient is male or female. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.8: Dynamic DiD effects on in-hospital mortality (emergency patients), by income deprivation of residential area (LSOA)



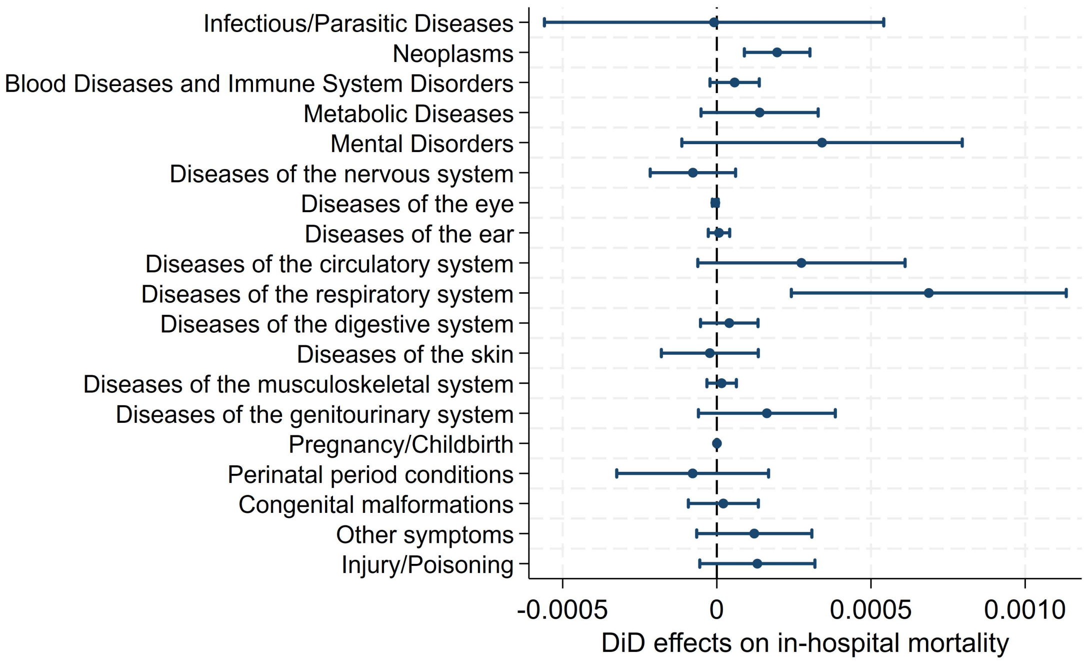
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating the triple-difference version of Equation 5.3 with the continuous exposure interacted with a dummy whether the patient belongs to the low- or high-deprivation group. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.9: DiD effects on in-hospital mortality (emergency patients), by ICD-10 diagnosis group



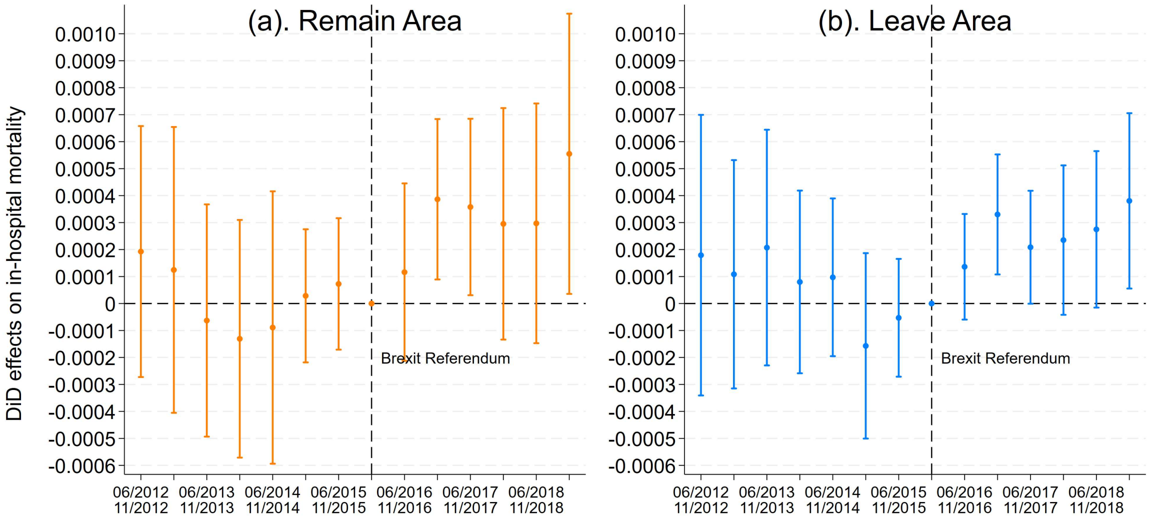
Note: This figure gives the estimated difference-in-difference effects of the Brexit referendum on patient-level health outcomes by diagnosis by estimating Equation 5.3 with the continuous exposure. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.10: DiD effects on in-hospital mortality (all patients), by ICD-10 diagnosis group



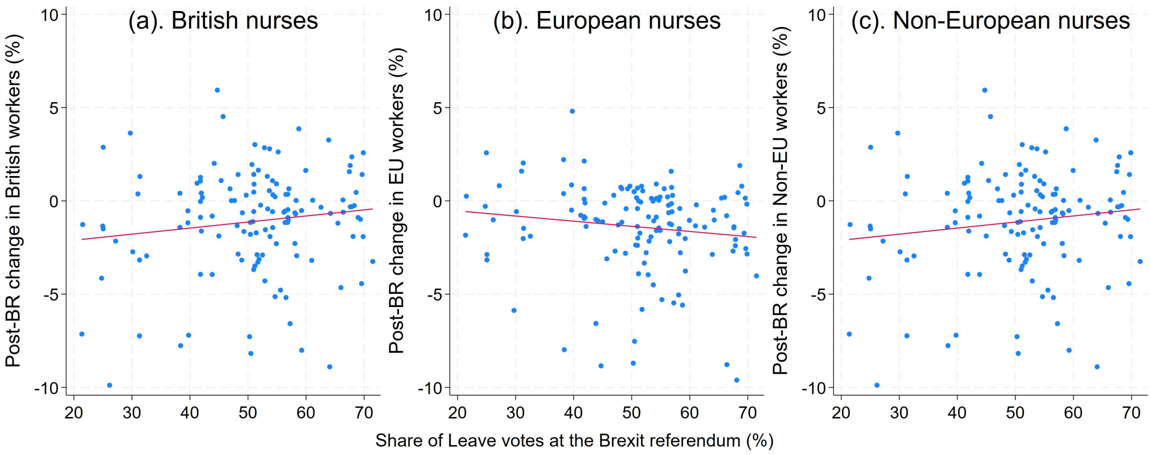
Note: This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by diagnosis by estimating Equation 5.3 with the continuous exposure. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.11: Heterogeneous DiD effects on in-hospital mortality, by prevalence of Brexit vote



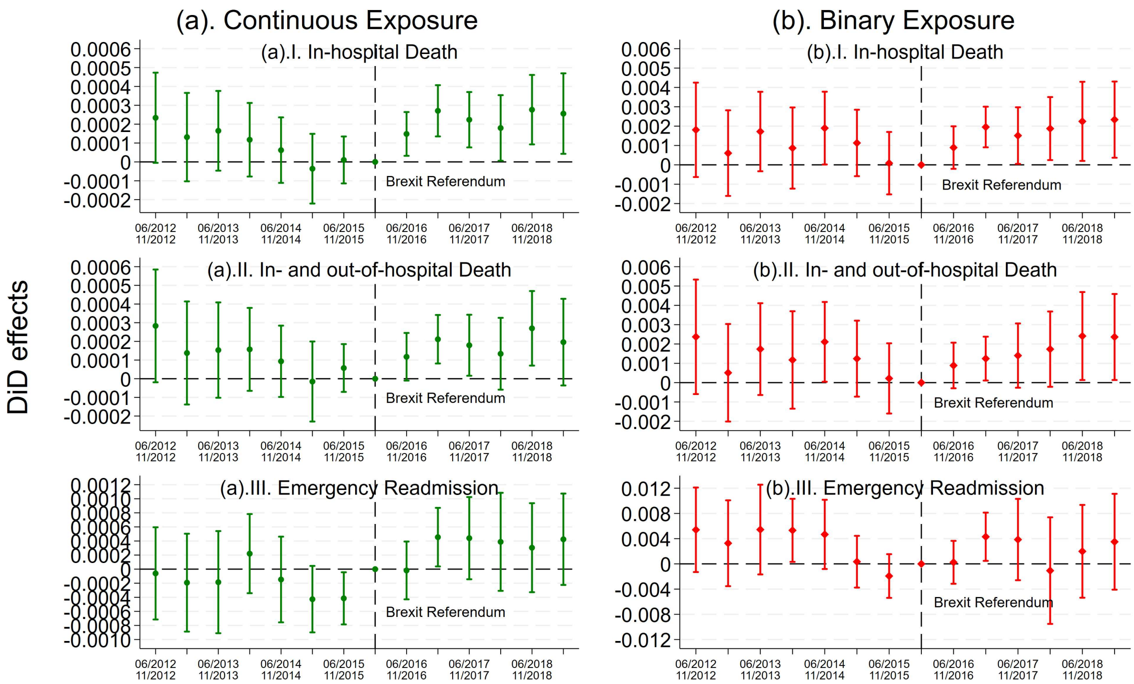
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating the triple-difference version of Equation 5.3 with the continuous exposure interacted with a dummy whether the provider is located in a 'Remain' or 'Leave' area. Hospital Organisations are allocated into the 'Remain' or 'Leave' groups based on whether the share of votes in support of leaving the EU in the June 2016 referendum in the postcode area of the hospital headquarter was respectively lower or higher than 50%. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.12: Share of Brexit 'Leave' votes and changes in the nurse workforce composition



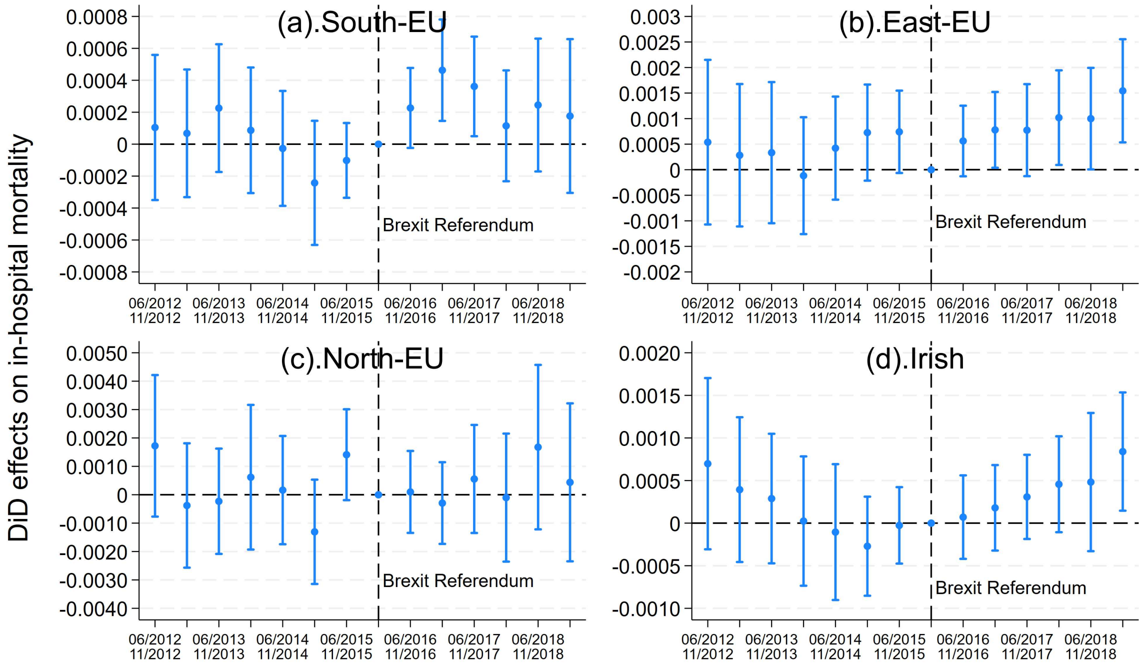
Note: This figures correlates the hospital provider-level post-referendum changes in the share of British, European, and non-European nurses with the local share of 'Leave' votes in the Brexit referendum.

Figure 5.B.13: Dynamic DiD effects on hospital care outcomes (emergency patients), treatment exposure based on period 06/2015-05/2016



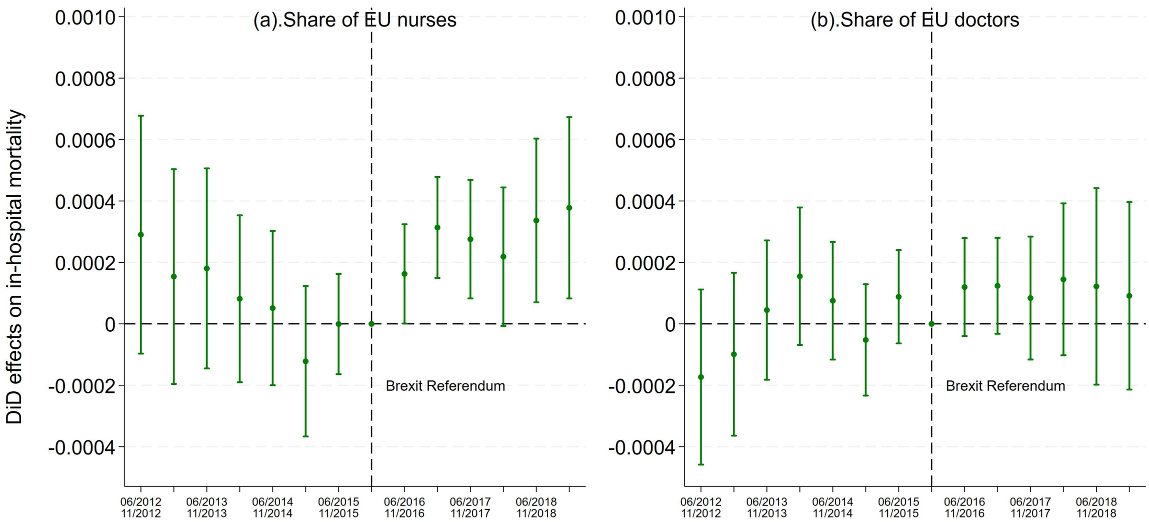
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by diagnosis by estimating Equation 5.3 with the continuous exposure. However, the continuous exposure is calculated based on the pre-policy year only in this figure. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.14: DiD effects on in-hospital mortality (emergency patients), by pre-referendum exposure variable measured by EU nurse nationality subgroups



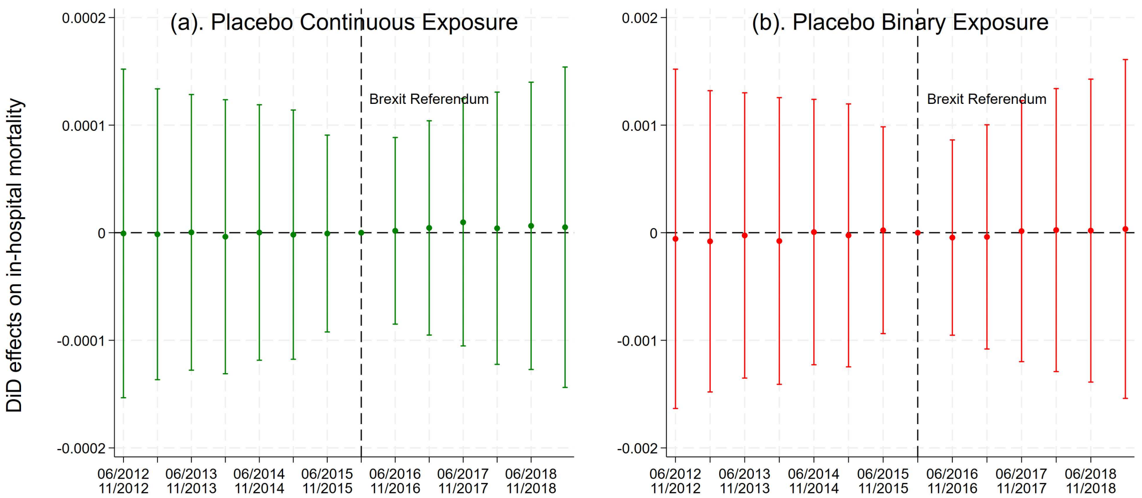
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by diagnosis by estimating Equation 5.3 with the continuous exposure. However, each panel refers to a different model where the continuous treatment exposure is calculated for the EU nationality subgroup specified in the panel header. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.15: DiD effects on in-hospital mortality (emergency patients), controlling for the share of EU senior doctors



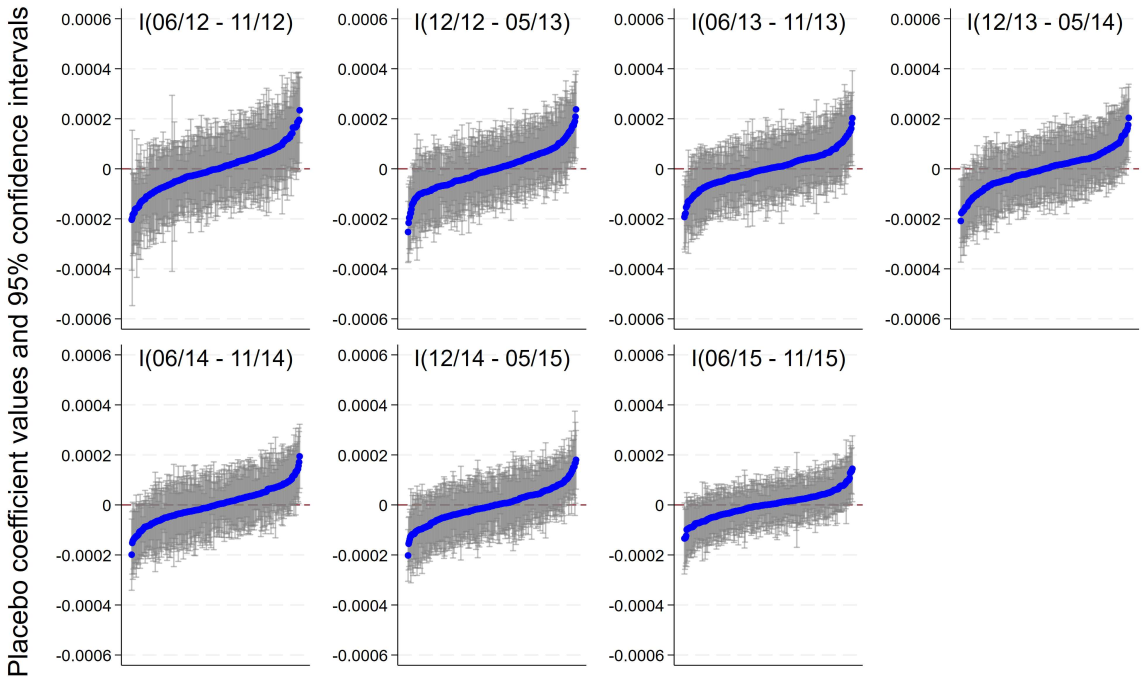
Note: N = 32,445,509 hospital emergency admissions. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by diagnosis by estimating Equation 5.3 with the continuous exposure. However, the continuous exposure is once calculated for nurses and once for doctors and both interactions are jointly estimated in the same regressions. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.16: DiD effects on in-hospital mortality (emergency patients), falsification tests based on randomized inference



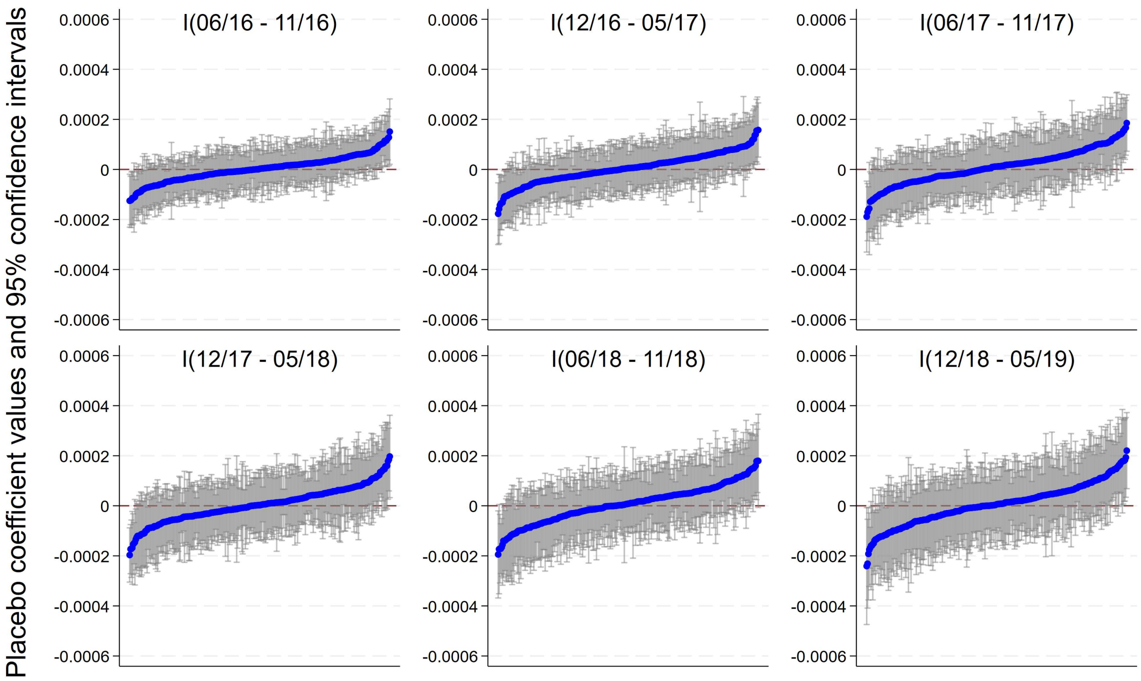
Note: This figure gives the dynamic treatment effects of the Brexit referendum on patient-level health outcomes by estimating Equation 5.3 with a continuous log-normally distributed placebo exposure (*panel a*) or a binary placebo exposure allocating hospital organisations into the top quartile of the exposure distribution at random (*panel b*).

Figure 5.B.17: Pre-treatment DiD effects on in-hospital mortality (emergency patients), falsification tests based on randomized inference



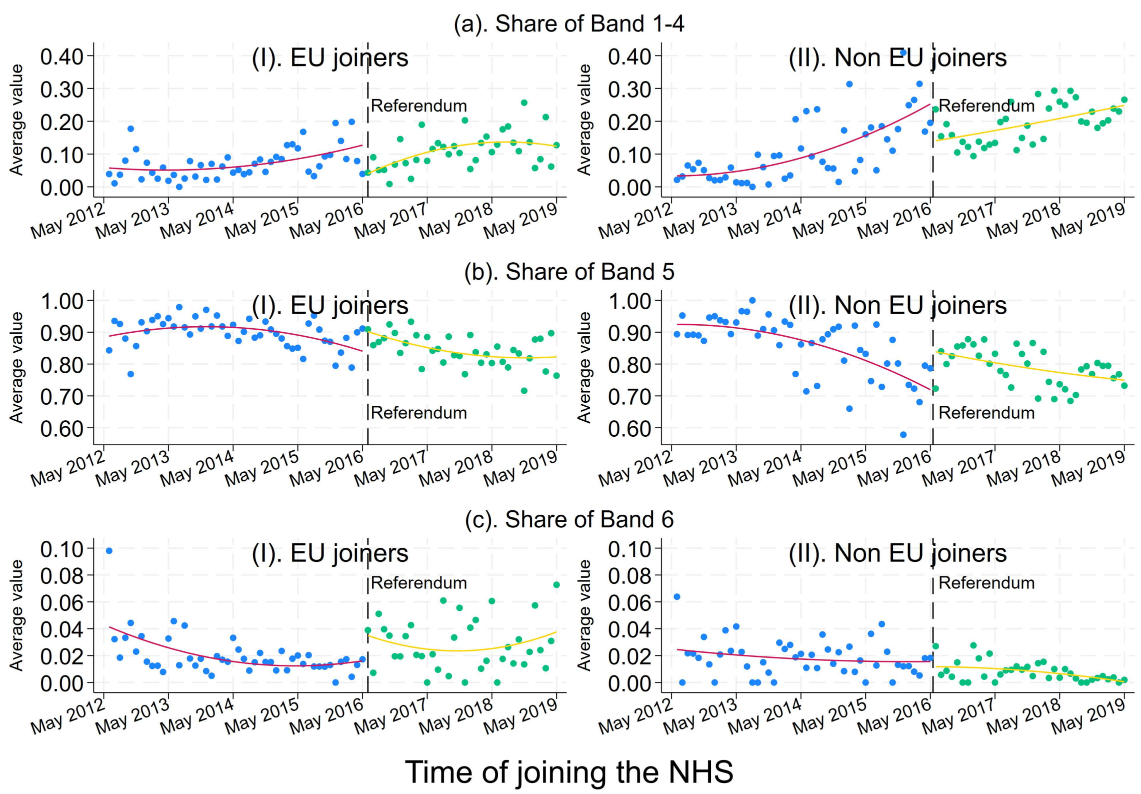
Note: This figure gives the dynamic pre-referendum effects of the Brexit referendum on emergency patients' in-hospital mortality by estimating Equation 5.3 with a continuous log-normally distributed placebo exposure variable (pre-referendum share of EU nurses employed at each hospital organization in the sample).

Figure 5.B.18: Post-treatment DiD effects on in-hospital mortality (emergency patients), falsification tests based on randomized inference



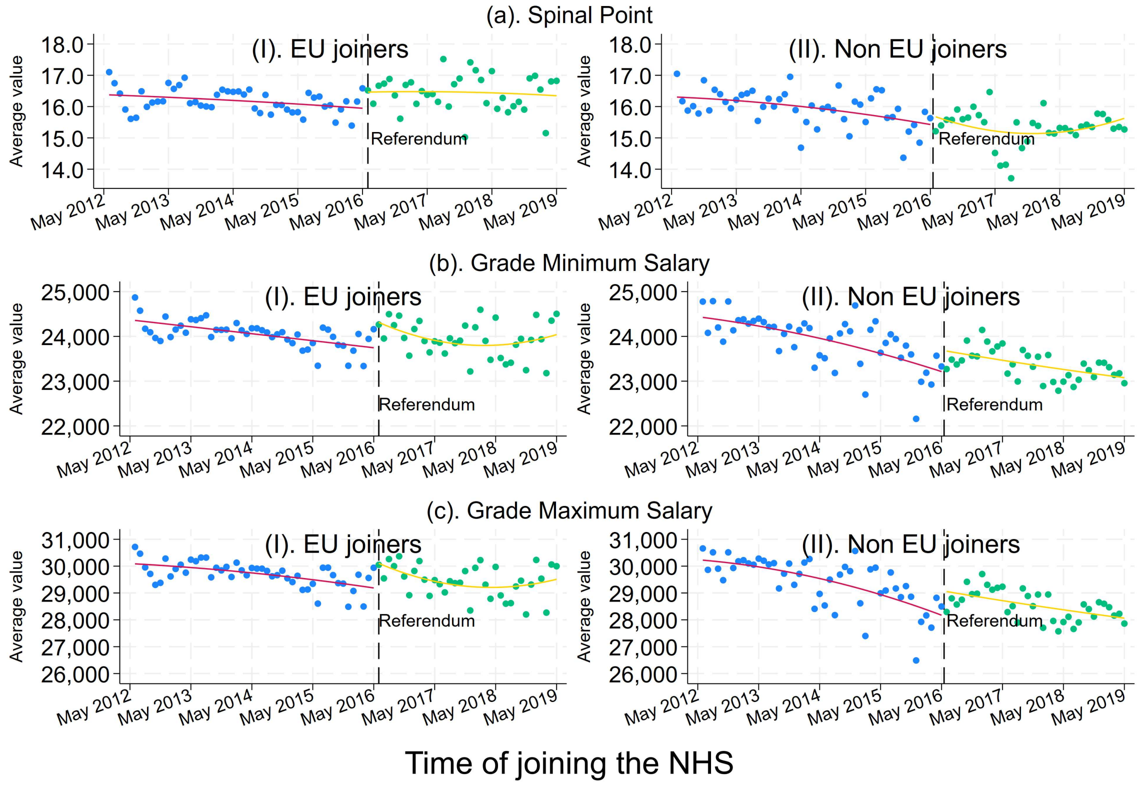
Note: This figure gives the dynamic post-referendum effects of the Brexit referendum on emergency patients' in-hospital mortality by estimating Equation 5.3 with a continuous log-normally distributed placebo exposure variable (pre-referendum share of EU nurses employed at each hospital organization in the sample).

Figure 5.B.19: Share of new NHS joiners by nationality subgroup and pay banding



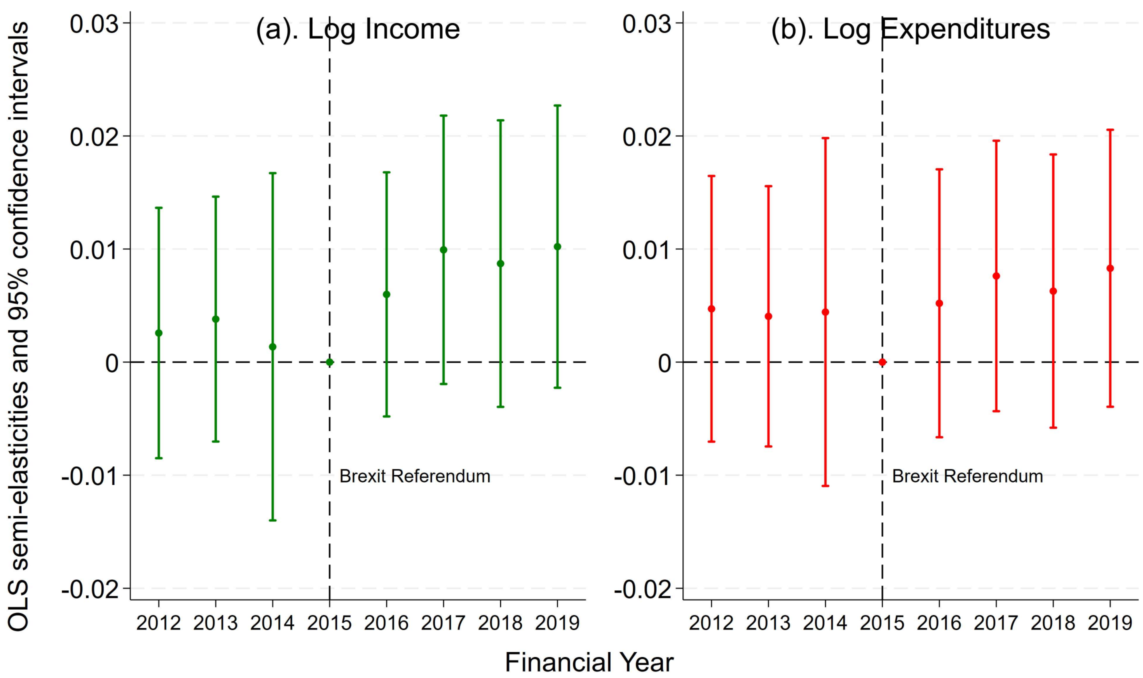
Note: This figure gives the share of new NHS joining nurses by nationality group (EU and non-EU nurses) and by wage band over time.

Figure 5.B.20: Share of new NHS joiners by nationality subgroup and salary grade



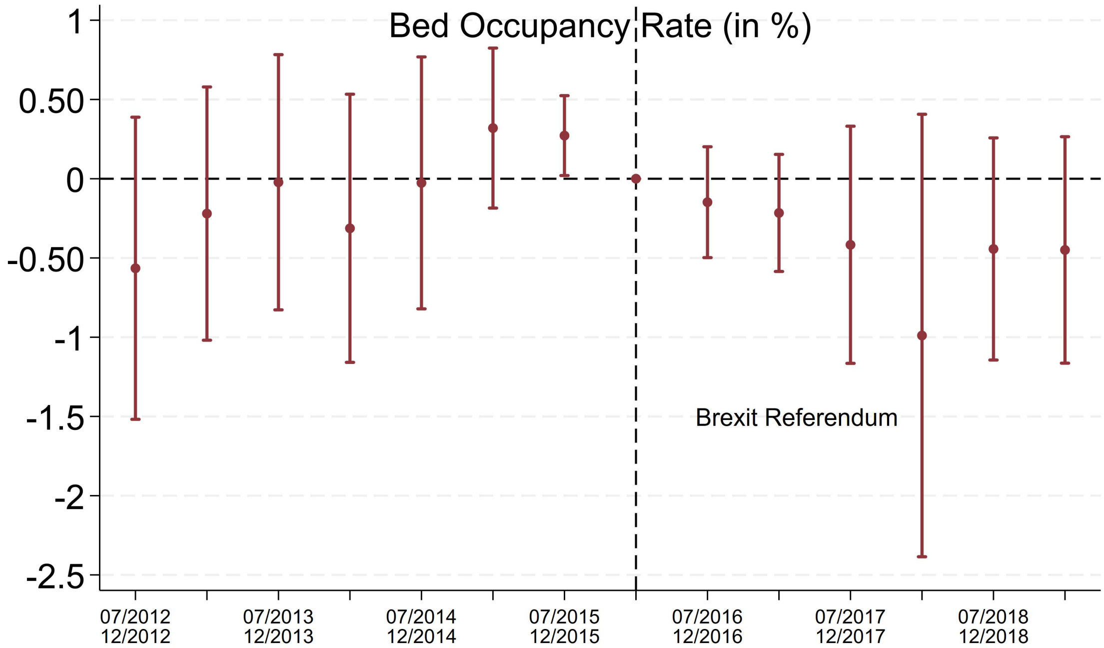
Note: This figure gives the share of new NHS joining nurses by nationality group (EU and non-EU nurses) and by salary grade over time.

Figure 5.B.21: Dynamic DiD effects on hospital financial positions



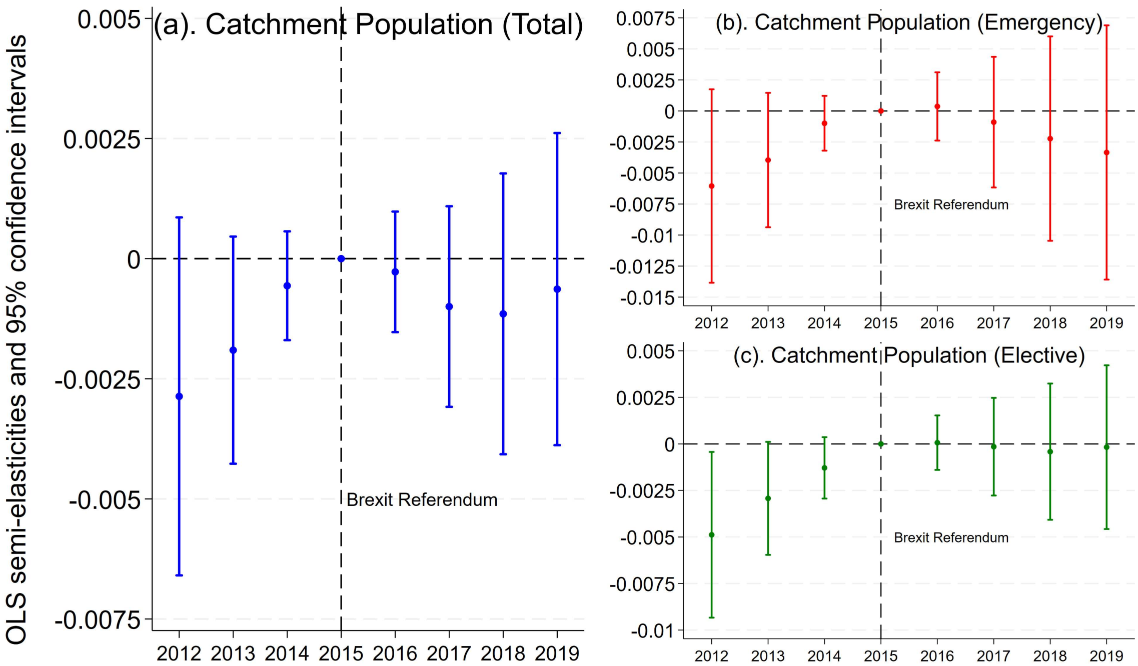
Note: N = 803 observations. $N_{clusters}$: 131 hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on provider-level financials by estimating Equation 5.3 with the continuous exposure at the hospital provider level. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.22: Dynamic DiD effects on hospital bed occupancy rates



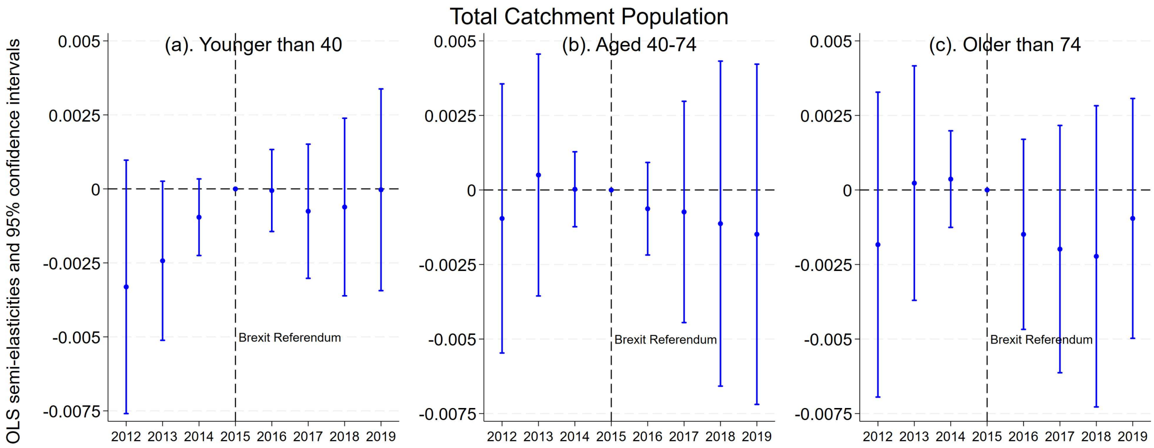
Note: This figure gives the dynamic treatment effects of the Brexit referendum on provider-level occupancy rates by estimating Equation 5.3 with the continuous exposure at the hospital provider level. Sample size: 3,569 hospital organization × quarter-year observations. $N_{clusters}$: 130 hospital organizations. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals. Joint Wald test $H_0 : \beta_{t_{-1}} = \beta_{t_{-2}} = ... = \beta_{t_{-k}} = 0$, $\{t_{-1}, t_{-2}, ..., t_{-k}\}$ < June 2016 (pre-referendum effects = 0): 2.71 (F-stat); 0.012 (p-value). Joint Wald test $H_0 : \beta_{t_1} = \beta_{t_2} = ... = \beta_{t_k} = 0$, $\{t_1, t_2, ..., t_k\} \geqslant$ June 2016 (post-referendum effects = 0): 1.21 (F-stat); 0.306 (p-value).

Figure 5.B.23: Dynamic DiD effects on patients' catchment population



Note: $N = 1{,}000$ observations. $N_{clusters} = 125$ hospital organizations. This figure gives the dynamic treatment effects of the Brexit referendum on the provider-level catchment area (total, emergency, and elective) by estimating Equation 5.3 with the continuous exposure at the hospital provider level. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Figure 5.B.24: Dynamic DiD effects on patients' catchment population by age group



Note: This figure gives the dynamic treatment effects of the Brexit referendum on the provider-level catchment area (by age group) by estimating Equation 5.3 with the continuous exposure at the hospital provider level. Robust standard errors clustered at hospital organization level; bands: 95% confidence intervals.

Table 5.B.1: Summary Statistics

	Pre-Brexit Referendum period		Post-Brexit Referendum period	
	Mean	Standard Deviation	Mean	Standard Deviation
Panel A: Patient health outcomes and covariates				
30-day in-hospital deaths per 100 admissions	3.26	(17.75)	3.04	(17.16)
30-day in- & out-of-hospital deaths per 100 admissions	4.23	(20.13)	4.05	(19.72)
30-day unplanned emergency readmissions per 100 admissions	15.20	(35.90)	16.08	(36.73)
Share of Female patients	0.52	(0.50)	0.53	(0.50)
Share of Income Deprivation Q1 (least deprived)	0.16	(0.36)	0.16	(0.37)
Share of Income Deprivation Q2	0.18	(0.39)	0.18	(0.39)
Share of Income Deprivation Q3	0.20	(0.40)	0.20	(0.40)
Share of Income Deprivation Q4	0.22	(0.41)	0.22	(0.41)
Share of Income Deprivation Q5 (most deprived)	0.24	(0.43)	0.24	(0.42)
Charlson comorbidities index (weighted)	5.07	(8.19)	5.76	(8.93)
Share of patients aged 0-18	0.15	(0.36)	0.15	(0.35)
Share of patients aged 18-29	0.10	(0.30)	0.09	(0.29)
Share of patients aged 30-39	0.08	(0.27)	0.08	(0.27)
Share of patients aged 40-49	0.09	(0.28)	0.08	(0.27)
Share of patients aged 50-59	0.10	(0.30)	0.10	(0.30)
Share of patients aged 60-69	0.12	(0.32)	0.11	(0.32)
Share of patients aged 70-79	0.15	(0.35)	0.15	(0.36)
Share of patients aged over 80	0.22	(0.42)	0.23	(0.42)
Share of Region: North West	0.15	(0.36)	0.15	(0.36)
Share of Region: North East	0.19	(0.39)	0.18	(0.39)
Share of Region: Midlands	0.16	(0.37)	0.16	(0.37)
Share of Region: East of England	0.10	(0.30)	0.10	(0.30)
Share of Region: South West	0.11	(0.31)	0.11	(0.32)
Share of Region: South East	0.16	(0.36)	0.17	(0.37)
Share of Region: London	0.13	(0.33)	0.13	(0.33)
Emergency Patients	9,471,534		8,287,621	
Emergency Admissions	17,610,090		14,835,419	
Panel B: Composition of hospital organization clinical workforce (monthly figures)				
Number of Employed Nurses	1,660.81	(939.77)	1,751.45	(1010.70)
Number of EU Employed Nurses	98.02	(116.58)	146.54	(166.17)
Number of Non-EU Employed Nurses	221.09	(248.54)	246.36	(278.90)
Number of Joiner Nurses	10.23	(11.91)	8.88	(11.51)
Number of EU Joiner Nurses	2.95	(6.04)	1.26	(2.36)
Number of non-EU Joiner Nurses	1.16	(2.74)	2.47	(4.80)
Number of Employed Senior Doctors	316.20	(186.86)	358.71	(211.21)
Number of EU Employed Senior Doctors	27.04	(19.34)	35.76	(27.25)
Number of non-EU Employed Senior Doctors	70.39	(46.82)	87.52	(54.26)
Number of Joiner Senior Doctors	1.72	(5.43)	1.50	(3.93)
Number of EU Joiner Senior Doctors	0.79	(1.17)	0.51	(1.06)
Number of non-EU Joiner Senior Doctors	0.73	(1.16)	1.13	(1.35)
Hospital-months records	6,288		6,288	
Number of hospital organizations	131		131	

Note: Panel A reports the descriptive statistics on patient-level health outcomes and covariates for all emergency patients admitted to the 131 NHS hospital organizations in the sample. *Panel B* reports monthly workforce composition figures at hospital organization level. Pre-Brexit referendum period: June 2012 to May 2016; post-Brexit referendum period: June 2016 to May 2019.

Table 5.B.2: Summary Statistics (all patients)

	Pre-Brexit Referendum period		Post-Brexit Referendum period	
	Mean	Standard Deviation	Mean	Standard Deviation
30-day in-hospital deaths per 100 admissions	1.21	(10.91)	1.15	(10.68)
30-day in- & out-of-hospital deaths per 100 admissions	1.62	(12.63)	1.59	(12.51)
30-day unplanned emergency readmissions per 100 admissions	7.80	(26.82)	8.38	(27.71)
Share of Female patients	0.55	(0.50)	0.55	(0.50)
Income Deprivation Q1 (least deprived)	0.18	(0.38)	0.18	(0.39)
Share of Income Deprivation Q2	0.19	(0.40)	0.20	(0.40)
Share of Income Deprivation Q3	0.20	(0.40)	0.20	(0.40)
Share of Income Deprivation Q4	0.21	(0.41)	0.21	(0.40)
Share of Income Deprivation Q5 (most deprived)	0.22	(0.41)	0.21	(0.41)
Charlson comorbidities Index (weighted)	3.26	(6.42)	3.88	(7.13)
Share of patients aged 0-18	0.10	(0.30)	0.10	(0.30)
Share of patients aged 18-29	0.10	(0.30)	0.09	(0.29)
Share of patients aged 30-39	0.10	(0.30)	0.10	(0.30)
Share of patients aged 40-49	0.10	(0.30)	0.09	(0.29)
Share of patients aged 50-59	0.13	(0.33)	0.13	(0.34)
Share of patients aged 60-69	0.16	(0.37)	0.15	(0.36)
Share of patients aged 70-79	0.17	(0.37)	0.18	(0.38)
Share of patients aged over 80	0.15	(0.35)	0.15	(0.36)
Share of Region: North West	0.15	(0.35)	0.14	(0.35)
Share of Region: North East	0.18	(0.38)	0.17	(0.38)
Share of Region: Midlands	0.15	(0.36)	0.15	(0.36)
Share of Region: East of England	0.10	(0.30)	0.10	(0.30)
Share of Region: South West	0.12	(0.32)	0.12	(0.32)
Share of Region: South East	0.15	(0.36)	0.16	(0.36)
Share of Region: London	0.15	(0.36)	0.15	(0.36)
Emergency and Non-Emergency Patients		18,124,652		15,723,616
Emergency and Non-Emergency Admissions		49,352,148		40,376,204

Note: Descriptive statistics on patient-level health outcomes and covariates for all emergency and elective patients admitted to the 131 NHS hospital organizations in the sample. Pre-Brexit referendum period: June 2012 to May 2016; post-Brexit referendum period: June 2016 to May 2019.

Table 5.B.3: Exposure to Brexit referendum shock and hospital patients' composition

	Emergency Patients				All Patients			
	Male	Age	Charlson Index	Income Depr. Index	Male	Age	Charlson Index	Income Depr. Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Continuous Exposure								
I(06/2016 - 05/2019) * Pre-BR share of EU nurses	-0.00004	-0.00461	-0.04224**	0.00009	0.00025	-0.01583	-0.03576***	0.00008
	(0.00030)	(0.05813)	(0.02038)	(0.00006)	(0.00034)	(0.03227)	(0.01291)	(0.00006)
Panel B: Binary Exposure								
I(06/2016 - 05/2019) * Pre-BR share of EU nurses	-0.00008	0.33757	-0.11230	0.00046	0.00171	0.08976	-0.16989	0.00065
	(0.00201)	(0.37725)	(0.17151)	(0.00086)	(0.00229)	(0.24016)	(0.11114)	(0.00074)
Observations (Hospital Admissions)	33,115,230	33,115,230	33,115,230	32,665,571	97,291,055	97,291,055	97,291,055	95,335,630

Note: This table presents the results of the difference-in-differences effects following Equation 5.3 with the continuous and binary exposure on different patient characteristics such as gender, age, severity, and deprivation for emergency only and all patients. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.4: Exposure to Brexit referendum shock and changes in hospital staff employment

	Δ Total Employment		Δ Total Joiners		Δ Total Leavers	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-BR share of EU nurses	0.0102	0.0007	0.0459	-0.8384	0.3582	1.0889
	(0.1758)	(0.3101)	(0.3347)	(0.6484)	(0.3938)	(0.8430)
Observations (Hospital Organizations)			131			
NHS region FE	No	Yes	No	Yes	No	Yes

Note: This table presents the results of the correlation of provider-level changes in total employment, total joiners, and total leavers with the continuous exposure share. The change in the total number of employees is computed between May 2019 and May 2016. The changes in the total number of joiners and leavers are computed as the changes in the cumulative number of joiners and leavers over the pre- and post-referendum periods. Robust standard errors are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.5: Exposure to Brexit referendum shock and changes in foreign nurses' employment shares

	Δ South-EU		Δ East EU		Δ North EU		Δ Irish	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: EU nurses								
Pre-BR share of EU nurses	-0.1646***	-0.2842***	0.0000	-0.0284	0.0033	-0.0022	-0.0515***	-0.0537*
	(0.0503)	(0.1061)	(0.0093)	(0.0211)	(0.0066)	(0.0128)	(0.0133)	(0.0274)
	Δ African		Δ Asian		Δ Other Non EU			
Panel B: Non-EU nurses								
Pre-BR share of EU nurses	-0.0144	0.0203	0.3428***	0.3338**	0.0096	0.0132		
	(0.0188)	(0.0356)	(0.0752)	(0.1348)	(0.0074)	(0.0162)		
Observations (Hospital Organizations)				131				
NHS region FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table presents the results of the correlation of provider-level changes in nurse employment by nationality group with the continuous exposure share. The changes in employment shares are computed between May 2019 and May 2016, namely the two endpoints of our post- and pre-BR analysis periods. Robust standard errors are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.6: Exposure to Brexit referendum shock and changes in foreign *joiner* nurses' employment shares

	Δ South-EU		Δ East EU		Δ North EU		Δ Irish	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: EU joining nurses								
Pre-BR share of EU nurses	-0.9977***	-1.6640***	0.0725	0.0460	0.0138	0.0224	-0.1474***	-0.2429**
	(0.2270)	(0.4634)	(0.0610)	(0.1502)	(0.0225)	(0.0458)	(0.0547)	(0.1159)
	Δ African		Δ Asian		Δ Other Non EU			
Panel B: Non-EU joining nurses								
Pre-BR share of EU nurses	0.0528	0.1546	1.0770***	0.9298*	0.0941**	0.1233		
	(0.0589)	(0.1288)	(0.2528)	(0.5002)	(0.0393)	(0.0890)		
Observations (Hospital Organizations)				131				
NHS region FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table presents the results of the correlation of provider-level changes in joining nurses by EU nationality group with the continuous exposure share. The changes in the number of joiners by nationality group are computed as the changes in the cumulative number of joiners by nationality over the pre- and post-referendum periods. Robust standard errors are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.7: Dynamic DiD estimates of the effects of nurse workforce exposure to Brexit on hospital care quality, emergency patients only

	Continuous Treatment			Binary Treatment		
	In-hospital death	In-and out-of-hospital death	Emergency Readmission	In-hospital death	In-and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)	(4)	(5)	(6)
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00027	0.00033	0.00006	0.00075	0.00157	0.00302
	(0.00020)	(0.00024)	(0.00054)	(0.00144)	(0.00173)	(0.00365)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00014	0.00013	0.00005	-0.00032	-0.00006	0.00356
	(0.00018)	(0.00020)	(0.00054)	(0.00127)	(0.00146)	(0.00382)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00019	0.00014	-0.00002	0.00072	0.00061	0.00226
	(0.00016)	(0.00020)	(0.00056)	(0.00122)	(0.00142)	(0.00387)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00010	0.00014	0.00050	-0.00010	0.00032	0.00216
	(0.00014)	(0.00016)	(0.00046)	(0.00103)	(0.00115)	(0.00278)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00006	0.00010	0.00003	0.00037	0.00078	0.00114
	(0.00013)	(0.00014)	(0.00045)	(0.00093)	(0.00101)	(0.00306)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00013	-0.00009	-0.00029	-0.00061	-0.00033	0.00003
	(0.00013)	(0.00015)	(0.00033)	(0.00094)	(0.00110)	(0.00234)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	0.00001	0.00011	-0.00040	-0.00028	0.00036	-0.00212
	(0.00008)	(0.00008)	(0.00027)	(0.00062)	(0.00066)	(0.00195)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00018**	0.00014	-0.00002	0.00057	0.00074	0.00159
	(0.00008)	(0.00009)	(0.00030)	(0.00062)	(0.00067)	(0.00205)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00033***	0.00023**	0.00064**	0.00121*	0.00101	0.00584***
	(0.00009)	(0.00009)	(0.00029)	(0.00066)	(0.00068)	(0.00215)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00029***	0.00023**	0.00074*	0.00122	0.00163*	0.00332
	(0.00010)	(0.00011)	(0.00044)	(0.00075)	(0.00085)	(0.00379)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00024**	0.00021	0.00076	0.00151*	0.00174*	0.00275
	(0.00011)	(0.00013)	(0.00050)	(0.00083)	(0.00102)	(0.00387)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00035***	0.00035**	0.00060	0.00187*	0.00224**	0.00178
	(0.00013)	(0.00014)	(0.00046)	(0.00103)	(0.00113)	(0.00394)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00039***	0.00035**	0.00084*	0.00246**	0.00263**	0.00217
	(0.00015)	(0.00016)	(0.00044)	(0.00111)	(0.00121)	(0.00421)
H_0: Pre-referendum coefficients = 0						
F-stat	1.123	1.350	1.991	0.88402	0.99710	1.08546
P-value	0.353	0.232	0.061	0.52113	0.43639	0.37633
H_0: Post-referendum coefficients = 0						
F-stat	2.823	2.002	2.659	0.980	0.979	3.020
P-value	0.013	0.070	0.018	0.441	0.442	0.008
Observations (emergency hospital admissions)	32,445,509					

Note: This table gives the pooled and dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.8: Dynamic DiD estimates of the effects of nurse workforce exposure to Brexit on hospital care quality, all patients

	Continuous Treatment			Binary Treatment		
	In-hospital death	In- and out-of-hospital death	Emergency Readmission	In-hospital death	In- and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Event-study						
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00013*	0.00017**	-0.00004	0.00054	0.00096	0.00069
	(0.00007)	(0.00008)	(0.00025)	(0.00053)	(0.00064)	(0.00181)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00004	0.00004	-0.00005	-0.00000	0.00012	0.00161
	(0.00006)	(0.00007)	(0.00023)	(0.00046)	(0.00054)	(0.00168)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00007	0.00006	-0.00003	0.00047	0.00047	0.00122
	(0.00006)	(0.00007)	(0.00023)	(0.00043)	(0.00049)	(0.00178)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00001	0.00001	0.00013	-0.00020	0.00002	0.00145
	(0.00005)	(0.00006)	(0.00021)	(0.00037)	(0.00041)	(0.00150)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00003	0.00004	-0.00002	0.00019	0.00023	0.00051
	(0.00005)	(0.00006)	(0.00017)	(0.00040)	(0.00043)	(0.00130)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00004	-0.00004	-0.00013	-0.00023	-0.00020	-0.00002
	(0.00004)	(0.00005)	(0.00012)	(0.00034)	(0.00041)	(0.00097)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	0.00002	0.00006**	-0.00014	-0.00000	0.00024	-0.00037
	(0.00003)	(0.00003)	(0.00012)	(0.00023)	(0.00023)	(0.00088)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00008***	0.00006*	0.00005	0.00032	0.00035	0.00054
	(0.00003)	(0.00003)	(0.00013)	(0.00024)	(0.00028)	(0.00096)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00012***	0.00010***	0.00028*	0.00067***	0.00071***	0.00214*
	(0.00003)	(0.00003)	(0.00014)	(0.00023)	(0.00026)	(0.00118)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00015***	0.00016***	0.00033*	0.00089***	0.00123***	0.00158
	(0.00004)	(0.00005)	(0.00020)	(0.00033)	(0.00039)	(0.00167)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00012**	0.00013**	0.00036*	0.00085**	0.00105**	0.00218
	(0.00005)	(0.00006)	(0.00021)	(0.00037)	(0.00046)	(0.00180)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00018***	0.00021***	0.00032	0.00097**	0.00128**	0.00179
	(0.00006)	(0.00007)	(0.00023)	(0.00045)	(0.00052)	(0.00196)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00020***	0.00022***	0.00038	0.00116**	0.00140**	0.00150
	(0.00007)	(0.00007)	(0.00025)	(0.00054)	(0.00063)	(0.00227)
H_0: Pre-referendum coefficients = 0						
F-stat	2.546	3.026	0.926	0.939	1.437	0.603
p-value	0.017	0.006	0.489	0.479	0.196	0.753
H_0: Post-referendum coefficients = 0						
F-stat	3.433	2.632	1.028	2.043	2.643	2.002
p-value	0.004	0.019	0.410	0.064	0.019	0.070
Observations (emergency & non-emergency hospital admissions)	89,728,352					

Note: This table gives the pooled and dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable - for all patients. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.9: Heterogeneous DiD effects by diagnosis group (emergency patients)

	Continuous Treatment					
	In-hospital death		In- and out-of-hospital death		Emergency Readmission	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
ICD chapter:						
Infectious/Parasitic Diseases	0.00000	(0.00033)	-0.00010	(0.00034)	0.00089*	(0.00047)
Neoplasms	0.00222***	(0.00077)	0.00189**	(0.00079)	0.00137	(0.00106)
Blood Diseases and Immune System Disorders	0.00017	(0.00021)	0.00048*	(0.00027)	-0.00021	(0.00186)
Metabolic Diseases	0.00021	(0.00023)	0.00004	(0.00024)	0.00214**	(0.00099)
Mental Disorders	0.00038	(0.00025)	0.00025	(0.00034)	0.00069	(0.00086)
Diseases of the nervous system	-0.00020	(0.00023)	-0.00023	(0.00027)	0.00072	(0.00067)
Diseases of the eye	-0.00026**	(0.00010)	-0.00029*	(0.00015)	-0.00073	(0.00097)
Diseases of the ear	0.00002	(0.00007)	0.00006	(0.00010)	0.00137	(0.00100)
Diseases of the circulatory system	0.00031	(0.00029)	0.00022	(0.00030)	0.00058	(0.00057)
Diseases of the respiratory system	0.00060**	(0.00030)	0.00045	(0.00032)	0.00066	(0.00051)
Diseases of the digestive system	0.00004	(0.00014)	0.00006	(0.00015)	0.00129**	(0.00058)
Diseases of the skin	-0.00005	(0.00014)	-0.00004	(0.00014)	-0.00031	(0.00134)
Diseases of the musculoskeletal system	-0.00003	(0.00010)	-0.00003	(0.00011)	-0.00031	(0.00076)
Diseases of the genitourinary system	0.00037*	(0.00020)	0.00015	(0.00022)	0.00104**	(0.00051)
Pregnancy/Childbirth	0.00000	(0.00001)	0.00001	(0.00001)	-0.00124	(0.00215)
Perinatal period conditions	-0.00001	(0.00006)	-0.00000	(0.00007)	0.00103	(0.00120)
Congenital malformations	-0.00007	(0.00036)	-0.00008	(0.00040)	0.00178	(0.00214)
Other symptoms	0.00017	(0.00014)	0.00015	(0.00016)	0.00038	(0.00060)
Injury/Poisoning	0.00015	(0.00012)	0.00011	(0.00012)	0.00083	(0.00052)

Note: This table gives the pooled regression results of Equation 5.3 for the continuous exposure for emergency patients only by diagnosis. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.10: Effects of nurse workforce exposure to Brexit on hospital care quality, excluding London hospitals from the sample

	Continuous Treatment			Binary Treatment		
	In-hospital death	In- and out-of-hospital death	Emergency Readmission	In-hospital death	In- and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)	(4)	(5)	(6)
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00033	0.00039	-0.00010	0.00160	0.00256	0.00236
	(0.00022)	(0.00026)	(0.00058)	(0.00149)	(0.00183)	(0.00402)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00019	0.00018	-0.00015	0.00024	0.00060	0.00263
	(0.00019)	(0.00022)	(0.00057)	(0.00138)	(0.00160)	(0.00422)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00023	0.00018	-0.00028	0.00122	0.00129	0.00122
	(0.00018)	(0.00022)	(0.00059)	(0.00133)	(0.00152)	(0.00417)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00019	0.00021	0.00033	0.00049	0.00086	0.00161
	(0.00014)	(0.00017)	(0.00048)	(0.00108)	(0.00123)	(0.00307)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00012	0.00017	-0.00013	0.00081	0.00128	-0.00001
	(0.00013)	(0.00015)	(0.00045)	(0.00101)	(0.00109)	(0.00328)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00014	-0.00009	-0.00047	-0.00045	-0.00012	-0.00172
	(0.00014)	(0.00017)	(0.00032)	(0.00105)	(0.00124)	(0.00234)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	-0.00002	0.00009	-0.00050*	-0.00018	0.00068	-0.00366*
	(0.00009)	(0.00009)	(0.00028)	(0.00069)	(0.00071)	(0.00195)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00020**	0.00016*	-0.00008	0.00093	0.00113	0.00079
	(0.00009)	(0.00009)	(0.00033)	(0.00066)	(0.00071)	(0.00231)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00038***	0.00028***	0.00060*	0.00182***	0.00174***	0.00516**
	(0.00009)	(0.00009)	(0.00032)	(0.00063)	(0.00058)	(0.00242)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00030***	0.00025**	0.00066	0.00160*	0.00203**	0.00216
	(0.00010)	(0.00012)	(0.00049)	(0.00081)	(0.00094)	(0.00428)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00028**	0.00025*	0.00074	0.00181*	0.00205*	0.00212
	(0.00012)	(0.00014)	(0.00056)	(0.00092)	(0.00114)	(0.00436)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00040***	0.00038***	0.00063	0.00253**	0.00277**	0.00111
	(0.00014)	(0.00014)	(0.00052)	(0.00108)	(0.00124)	(0.00448)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00044***	0.00040**	0.00083*	0.00304**	0.00317**	0.00113
	(0.00015)	(0.00016)	(0.00049)	(0.00118)	(0.00131)	(0.00477)
Observations (emergency hospital admissions)	28,357,810					

Note: This table gives the pooled and dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable - excluding patients and hospitals from London. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.11: Effects of nurse workforce exposure to Brexit on hospital care quality, using a Brexit shock exposure computed over 06/2015-05/2016

	Continuous Treatment			Binary Treatment		
	In-hospital death	In- and out-of-hospital death	Emergency Readmission	In-hospital death	In- and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)	(4)	(5)	(6)
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00023*	0.00028*	-0.00002	0.00122	0.00197	0.00546*
	(0.00012)	(0.00015)	(0.00032)	(0.00120)	(0.00137)	(0.00288)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00013	0.00013	-0.00017	-0.00012	-0.00012	0.00234
	(0.00012)	(0.00014)	(0.00035)	(0.00106)	(0.00119)	(0.00329)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00016	0.00015	-0.00017	0.00084	0.00098	0.00397
	(0.00011)	(0.00013)	(0.00037)	(0.00102)	(0.00111)	(0.00356)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00012	0.00016	0.00024	-0.00002	0.00027	0.00332
	(0.00010)	(0.00011)	(0.00029)	(0.00113)	(0.00131)	(0.00267)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00005	0.00009	-0.00012	0.00110	0.00143	0.00307
	(0.00009)	(0.00010)	(0.00032)	(0.00100)	(0.00107)	(0.00278)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00005	-0.00003	-0.00041*	-0.00002	0.00020	-0.00115
	(0.00009)	(0.00011)	(0.00025)	(0.00095)	(0.00107)	(0.00218)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	0.00000	0.00005	-0.00042**	-0.00051	-0.00021	-0.00229
	(0.00006)	(0.00007)	(0.00019)	(0.00087)	(0.00097)	(0.00179)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00014**	0.00011*	-0.00002	0.00026	0.00042	0.00090
	(0.00006)	(0.00006)	(0.00021)	(0.00061)	(0.00062)	(0.00187)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00027***	0.00022***	0.00045**	0.00159***	0.00126**	0.00534**
	(0.00007)	(0.00007)	(0.00022)	(0.00058)	(0.00061)	(0.00214)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00022***	0.00018**	0.00043	0.00127	0.00154*	0.00403
	(0.00007)	(0.00008)	(0.00030)	(0.00082)	(0.00089)	(0.00365)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00017*	0.00011	0.00037	0.00128	0.00115	0.00116
	(0.00009)	(0.00010)	(0.00036)	(0.00090)	(0.00104)	(0.00483)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00028***	0.00027***	0.00030	0.00213*	0.00260**	0.00010
	(0.00009)	(0.00010)	(0.00032)	(0.00112)	(0.00122)	(0.00414)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00024**	0.00017	0.00042	0.00214**	0.00228**	0.00023
	(0.00011)	(0.00012)	(0.00033)	(0.00105)	(0.00112)	(0.00430)
Observations (emergency hospital admissions)	32,445,509					

Note: This table gives the pooled and dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable and the binary exposure variable where both variables are calculated based on the last pre-policy year only. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.12: Effects of nurse workforce exposure to Brexit on hospital care quality, including the pre Brexit referendum share of non-EU nurses as control

	Continuous Treatment		
	In-hospital death	In- and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00026	0.00030	0.00008
	(0.00019)	(0.00023)	(0.00054)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00014	0.00010	0.00009
	(0.00017)	(0.00019)	(0.00054)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00020	0.00014	0.00001
	(0.00016)	(0.00019)	(0.00055)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00011	0.00013	0.00050
	(0.00013)	(0.00015)	(0.00045)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00007	0.00008	0.00004
	(0.00013)	(0.00014)	(0.00045)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00011	-0.00009	-0.00029
	(0.00013)	(0.00015)	(0.00033)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	-0.00000	0.00009	-0.00042
	(0.00008)	(0.00008)	(0.00026)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00016*	0.00013	-0.00011
	(0.00008)	(0.00010)	(0.00030)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00033***	0.00023**	0.00057*
	(0.00009)	(0.00009)	(0.00029)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00025***	0.00019*	0.00068
	(0.00009)	(0.00011)	(0.00044)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00022*	0.00018	0.00067
	(0.00012)	(0.00014)	(0.00050)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00029**	0.00029**	0.00054
	(0.00013)	(0.00014)	(0.00047)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00034**	0.00030*	0.00087*
	(0.00015)	(0.00016)	(0.00045)
I(06/2012 - 11/2012) * Pre-BR share of Non-EU nurses	0.00001	0.00008	-0.00004
	(0.00010)	(0.00011)	(0.00021)
I(12/2012 - 05/2013) * Pre-BR share of Non-EU nurses	-0.00001	0.00008	-0.00012
	(0.00008)	(0.00010)	(0.00022)
I(06/2013 - 11/2013) * Pre-BR share of Non-EU nurses	-0.00004	0.00002	-0.00008
	(0.00006)	(0.00008)	(0.00024)
I(12/2013 - 05/2014) * Pre-BR share of Non-EU nurses	-0.00003	0.00001	0.00001
	(0.00007)	(0.00008)	(0.00020)
I(06/2014 - 11/2014) * Pre-BR share of Non-EU nurses	-0.00001	0.00005	-0.00002
	(0.00007)	(0.00007)	(0.00020)
I(12/2014 - 05/2015) * Pre-BR share of Non-EU nurses	-0.00006	0.00001	0.00000
	(0.00006)	(0.00007)	(0.00018)
I(06/2015 - 11/2015) * Pre-BR share of Non-EU nurses	0.00004	0.00006	0.00007
	(0.00004)	(0.00005)	(0.00013)
I(06/2016 - 11/2016) * Pre-BR share of Non-EU nurses	0.00004	0.00004	0.00023**
	(0.00005)	(0.00005)	(0.00010)
I(12/2016 - 05/2017) * Pre-BR share of Non-EU nurses	-0.00001	0.00001	0.00018
	(0.00005)	(0.00005)	(0.00013)
I(06/2017 - 11/2017) * Pre-BR share of Non-EU nurses	0.00010*	0.00012*	0.00016
	(0.00005)	(0.00006)	(0.00021)
I(12/2017 - 05/2018) * Pre-BR share of Non-EU nurses	0.00005	0.00007	0.00024
	(0.00006)	(0.00007)	(0.00029)
I(06/2018 - 11/2018) * Pre-BR share of Non-EU nurses	0.00018***	0.00019**	0.00016
	(0.00007)	(0.00008)	(0.00025)
I(12/2018 - 05/2019) * Pre-BR share of Non-EU nurses	0.00014*	0.00014	-0.00009
	(0.00008)	(0.00009)	(0.00025)
Observations (emergency hospital admissions)		32,445,509	

Note: This table gives the pooled and dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable for in-hospitals deaths, in- and out-of-hospital deaths, and emergency readmissions. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.13: Effects of nurse workforce exposure to Brexit on hospital care quality, including the pre Brexit referendum share of EU doctors as control

	Continuous Treatment					
	In-hospital death	In- and out-of-hospital death	Emergency Readmission	In-hospital death	In- and out-of-hospital death	Emergency Readmission
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Event study						
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	0.00028	0.00034	0.00007			
	(0.00019)	(0.00024)	(0.00054)			
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	0.00015	0.00015	0.00000			
	(0.00017)	(0.00020)	(0.00054)			
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	0.00018	0.00015	-0.00003			
	(0.00016)	(0.00019)	(0.00054)			
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	0.00007	0.00010	0.00048			
	(0.00013)	(0.00015)	(0.00045)			
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	0.00005	0.00009	0.00003			
	(0.00012)	(0.00014)	(0.00043)			
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00011	-0.00008	-0.00033			
	(0.00012)	(0.00015)	(0.00031)			
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	-0.00000	0.00009	-0.00040			
	(0.00008)	(0.00008)	(0.00026)			
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	0.00016**	0.00013	0.00003			
	(0.00008)	(0.00009)	(0.00027)			
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	0.00030***	0.00021**	0.00068**			
	(0.00008)	(0.00008)	(0.00027)			
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	0.00027***	0.00022**	0.00080*			
	(0.00009)	(0.00011)	(0.00042)			
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	0.00022**	0.00018	0.00079			
	(0.00011)	(0.00013)	(0.00049)			
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	0.00033**	0.00033**	0.00069			
	(0.00013)	(0.00014)	(0.00045)			
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	0.00038***	0.00035**	0.00089**			
	(0.00014)	(0.00015)	(0.00042)			
I(06/2012 - 11/2012) * Pre-BR share of EU doctors	-0.00017	-0.00019	-0.00041	-0.00014	-0.00015	-0.00041
	(0.00014)	(0.00019)	(0.00040)	(0.00014)	(0.00018)	(0.00040)
I(12/2012 - 05/2013) * Pre-BR share of EU doctors	-0.00010	-0.00014	-0.00004	-0.00008	-0.00012	-0.00004
	(0.00013)	(0.00018)	(0.00040)	(0.00013)	(0.00017)	(0.00039)
I(06/2013 - 11/2013) * Pre-BR share of EU doctors	0.00005	-0.00002	-0.00022	0.00006	-0.00001	-0.00023
	(0.00011)	(0.00016)	(0.00036)	(0.00012)	(0.00016)	(0.00037)
I(12/2013 - 05/2014) * Pre-BR share of EU doctors	0.00016	0.00020	-0.00009	0.00016	0.00021	-0.00004
	(0.00011)	(0.00014)	(0.00033)	(0.00011)	(0.00014)	(0.00032)
I(06/2014 - 11/2014) * Pre-BR share of EU doctors	0.00008	0.00006	-0.00041	0.00008	0.00006	-0.00041
	(0.00010)	(0.00012)	(0.00034)	(0.00010)	(0.00012)	(0.00033)
I(12/2014 - 05/2015) * Pre-BR share of EU doctors	-0.00005	0.00000	0.00001	-0.00007	-0.00001	-0.00003
	(0.00009)	(0.00011)	(0.00027)	(0.00010)	(0.00011)	(0.00027)
I(06/2015 - 11/2015) * Pre-BR share of EU doctors	0.00009	0.00010	-0.00006	0.00009	0.00011	-0.00010
	(0.00008)	(0.00009)	(0.00021)	(0.00008)	(0.00009)	(0.00021)
I(06/2016 - 11/2016) * Pre-BR share of EU doctors	0.00012	0.00009	-0.00051*	0.00014*	0.00010	-0.00050*
	(0.00008)	(0.00010)	(0.00027)	(0.00008)	(0.00009)	(0.00027)
I(12/2016 - 05/2017) * Pre-BR share of EU doctors	0.00013	0.00013	-0.00044	0.00016*	0.00015	-0.00037
	(0.00008)	(0.00011)	(0.00031)	(0.00008)	(0.00011)	(0.00033)
I(06/2017 - 11/2017) * Pre-BR share of EU doctors	0.00009	0.00010	-0.00065	0.00011	0.00013	-0.00056
	(0.00010)	(0.00013)	(0.00051)	(0.00010)	(0.00012)	(0.00053)
I(12/2017 - 05/2018) * Pre-BR share of EU doctors	0.00015	0.00031**	-0.00019	0.00017	0.00032**	-0.00011
	(0.00013)	(0.00014)	(0.00054)	(0.00012)	(0.00013)	(0.00056)
I(06/2018 - 11/2018) * Pre-BR share of EU doctors	0.00012	0.00022	-0.00062	0.00016	0.00025*	-0.00054
	(0.00016)	(0.00015)	(0.00050)	(0.00016)	(0.00015)	(0.00053)
I(12/2018 - 05/2019) * Pre-BR share of EU doctors	0.00009	0.00016	-0.00046	0.00013	0.00020	-0.00038
	(0.00015)	(0.00016)	(0.00045)	(0.00015)	(0.00016)	(0.00049)
Observations (emergency hospital admissions)		32,445,509			32,445,509	

Note: This table gives the dynamic regression results of Equation 5.3 for all interactions with the continuous exposure variable for nurses and doctors separately. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.14: Effects of nurse workforce exposure to Brexit on (log) hospital admissions

	Continuous Treatment			Binary Treatment		
	log(Total) (1)	log(Emergency) (2)	log(Non-Emergency) (3)	log(Total) (4)	log(Emergency) (5)	log(Non-Emergency) (6)
I(06/2012 - 11/2012) * Pre-BR share of EU nurses	-0.00373	-0.00238	-0.00456	-0.05270	-0.04814	-0.05592
	(0.00382)	(0.00689)	(0.00424)	(0.03513)	(0.04723)	(0.03725)
I(12/2012 - 05/2013) * Pre-BR share of EU nurses	-0.00334	-0.00012	-0.00413	-0.04806	-0.02855	-0.05691*
	(0.00361)	(0.00670)	(0.00410)	(0.03216)	(0.04732)	(0.03369)
I(06/2013 - 11/2013) * Pre-BR share of EU nurses	-0.00137	0.00081	-0.00343	-0.02903	-0.00231	-0.04449
	(0.00302)	(0.00585)	(0.00314)	(0.03396)	(0.05096)	(0.03197)
I(12/2013 - 05/2014) * Pre-BR share of EU nurses	-0.00106	0.00191	-0.00311	-0.01650	-0.00054	-0.02959
	(0.00256)	(0.00473)	(0.00271)	(0.03292)	(0.04522)	(0.02986)
I(06/2014 - 11/2014) * Pre-BR share of EU nurses	-0.00064	0.00189	-0.00238	-0.01792	-0.00450	-0.03142
	(0.00214)	(0.00455)	(0.00207)	(0.03195)	(0.04271)	(0.02878)
I(12/2014 - 05/2015) * Pre-BR share of EU nurses	-0.00147	-0.00034	-0.00187	-0.00847	0.00553	-0.01593
	(0.00167)	(0.00282)	(0.00168)	(0.02091)	(0.02836)	(0.01904)
I(06/2015 - 11/2015) * Pre-BR share of EU nurses	0.00085	0.00076	0.00056	0.00944	0.01698	0.00435
	(0.00122)	(0.00198)	(0.00140)	(0.00938)	(0.01586)	(0.01002)
I(06/2016 - 11/2016) * Pre-BR share of EU nurses	-0.00092	-0.00354	-0.00124	0.00533	-0.00903	0.00866
	(0.00139)	(0.00308)	(0.00162)	(0.00952)	(0.01643)	(0.01069)
I(12/2016 - 05/2017) * Pre-BR share of EU nurses	-0.00158	-0.00038	-0.00195	-0.00624	0.00114	-0.00662
	(0.00128)	(0.00327)	(0.00160)	(0.00967)	(0.01524)	(0.01196)
I(06/2017 - 11/2017) * Pre-BR share of EU nurses	-0.00314*	-0.00577	-0.00370	-0.02036	-0.02932	-0.02270
	(0.00189)	(0.00493)	(0.00261)	(0.01303)	(0.02788)	(0.01658)
I(12/2017 - 05/2018) * Pre-BR share of EU nurses	-0.00315	-0.00355	-0.00326	-0.03140**	-0.05502*	-0.02233
	(0.00224)	(0.00440)	(0.00318)	(0.01530)	(0.02940)	(0.02087)
I(06/2018 - 11/2018) * Pre-BR share of EU nurses	-0.00120	-0.00228	-0.00256	-0.02291	-0.03381	-0.02461
	(0.00275)	(0.00458)	(0.00347)	(0.01923)	(0.03254)	(0.02342)
I(12/2018 - 05/2019) * Pre-BR share of EU nurses	-0.00278	-0.00335	-0.00404	-0.04087*	-0.03903	-0.04058
	(0.00321)	(0.00566)	(0.00383)	(0.02266)	(0.03561)	(0.02572)
Observations (hospital organisations x months)		11,004			11,004	

Note: This table gives the pooled regression results of Equation 5.3 for the continuous and binary exposure variable. The outcome is the natural logarithm of the number of total, emergency, or non-emergency patients. Robust standard errors clustered at hospital organization level. $N_{clusters}$: 131 acute care NHS hospital organizations. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5.B.15: Pay grades of foreign joiner nurses, before and after the Brexit referendum

	Pre-BR period		Post-BR period	
	Mean	St.Dev.	Mean	St.Dev.
Share of Band 1-4	0.081	0.054	0.166	0.059
Share of Band 5	0.887	0.052	0.808	0.052
Share of Band 6	0.019	0.012	0.014	0.010
Spinal Point	16.119	0.350	15.702	0.423
Grade Minimum Salary	24,012.54	322.218	23,542.52	340.501
Grade Maximum Salary	29,625.81	524.921	28,829.16	565.240

Note: Descriptive statistics of the EU and non-EU joiner nurse pay grade variables, split by the two sub-periods (June 2012 - May 2016; June 2016 - May 2019) in our sample, based on NHS hospitals ESR records.

Bibliography

Abowd, J. M., Kramarz, F. and Margolis, D. N. (1999), 'High Wage Workers and High Wage Firms', *Econometrica* **67**(2), 251–333.

Adda, J., Dustmann, C. and Görlach, J.-S. (2022), 'The Dynamics of Return Migration, Human Capital Accumulation, and Wage Assimilation', *Review of Economic Studies* **89**(6), 2841–2871.

Alabrese, E., Becker, S. O., Fetzer, T. and Novy, D. (2019), 'Who Voted for Brexit? Individual and Regional Data Combined', *European Journal of Political Economy* **56**, 132–150.

Becker, S. O., Fetzer, T. and Novy, D. (2017), 'Who Voted for Brexit? A Comprehensive District-Level Analysis', *Economic Policy* **32**(92), 601–650.

Beesack, P. R. (1957), 'A Note on an Integral Inequality', *Proceedings of the American Mathematical Society* **8**(5), 875–879.

Bertheau, A., Cahuc, P., Jäger, S. and Vejlin, R. (2022), 'Turnover Costs: Evidence from Unexpected Worker Seperations'. Working Paper.

Billings, J., Blunt, I., Steventon, A., Georghiou, T., Lewis, G. and Bardsley, M. (2012), 'Development of a Predictive Model to Identify Inpatients at Risk of Re-admission Within 30 Days of Discharge (PARR-30)', *BMJ Open* **2**(4), e001667.

Borjas, G. J. (2003), 'The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market', *The Quarterly Journal of Economics* **118**(4), 1335–1374.

Borjas, G. J. and Bratsberg, B. (1996), 'Who Leaves? The Outmigration of the Foreign-Born', *Review of Economics and Statistics* **78**(1), 165–176.

Born, B., , Müller, G. J., Schularick, M. and Sedlacek, P. (2019), 'The Costs of Economic Nationalism: Evidence from the Brexit Experiment', *Economic Journal* **129**(10), 2722–2744.

Breinlich, H., Leromain, E., Novy, D. and Sampson, T. (2020), 'Voting with Their Money: Brexit and Outward Investment by UK Firms', *European Economic Review* **124**, 103400.

Breinlich, H., Leromain, E., Novy, D. and Sampson, T. (2022), 'The Brexit Vote, Inflation and U.K. Living Standards"', *International Economic Review* **63**(1), 63–93.

Brenoe, A. A., Canaan, S., Harmon, N. and Royer, H. (2024), 'Is Parental Leave Costly for Firms and Coworkers?', *Journal of Labor Economics* **42**(4), 1135–1174.

Callaway, B., Goodman-Bacon, A. and Sant'Anna, P. H. (2024), Difference-in-Differences with a Continuous Treatment, Technical report, National Bureau of Economic Research.

Card, D. (2001), 'Immigrant Inflows, Native Outflows, and the Local Labour Market Impact of Higher Immigration', *Journal of Labor Economics* **19**(1), 22–64.

Chan Jr., D. C. and Chen, Y. (2023), 'The Productivity of Professions: Evidence from the Emergency Department'. Working Paper.

Charlson, M. E., Pompei, P., Ales, K. L. and MacKenzie, C. R. (1987), 'A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation', *Journal of Chronic Diseases* **40**(5), 373–383.

Choudhury, P., Doran, K., Marinoni, A. and Yoon, C. (2022), 'Loss of Peers and Individual Worker Performance: Evidence from H-1B Visa Denials', *CESifo WP* .

Cooper, Z., Gibbons, S., Jones, S. and McGuire, A. (2011), 'Does Hospital Competition Save Lives? Evidence from the English NHS Patient Choice Reforms', *The Economic Journal* **121**(554), F228–F260.

Cortes, P. and Pan, J. (2015), 'The Relative Quality of Foreign-Educated Nurses in the United States', *Journal of Human Resources* **50**(4), 1009–1050.

Cutler, D. M. and McClellan, M. (2001), 'Is Technological Change in Medicine Worth it?', *Health affairs* **20**(5), 11–29.

Davies, R. B. and Studnicka, Z. (2018), 'The Heterogeneous Impact of Brexit: Early Evidence from the FTSE', *European Economic Review* **110**, 1–17.

Doran, K., Gelber, A. and Isen, A. (2022), 'The Effects of High-Skilled Immigration Policy on Firms: Evidence from Visa Lotteries', *Journal of Political Economy* **130**(10), 2501–2533.

Dustmann, C., Fabbri, F. and Preston, I. (2005), 'The Impact of Immigration on the British Labour Market', *The Economic Journal* **115**(507), F324–F341.

Dustmann, C. and Frattini, T. (2014), 'The Fiscal Effects of Immigration to the UK', *The Economic Journal* **124**(580), F593–F643.

Dustmann, C., Frattini, T. and Preston, I. (2013), 'The Effect of Immigration Along the Distribution of Wages', *Review of Economic Studies* **132**(1), 435–483.

Dustmann, C. and Görlach, J.-S. (2016), 'The Economics of Temporary Migrations', *Journal of Economic Literature* **54**(1), 98–136.

European Commission (2020), 'Analysis of Shortage and Surplus Occupations'.
URL: *https://op.europa.eu/en/publication-detail/-/publication/22189434-395d-11eb-b27b-01aa75ed71a1/language-en*

European Labour Authority (2023), 'Report on Labour Shortages and Surpluses'.
URL: *https://www.ela.europa.eu/sites/default/files/2023-03/eures-labour-shortages-report-2022.pdf*

European Parliament (2023), 'Free Movement of Persons'. `https://www.europarl.europa.eu/factsheets/en/sheet/147/free-movement-of-persons`.

Eurostat (2023*a*), 'Job Vacancy and Unemployment Rates - Beveridge Curve'.
URL: *https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Job_vacancy_and_unemployment_rates_-_Beveridge_curve*

Eurostat (2023*b*), 'Population Projections in the EU'.
URL: *https://ec.europa.eu/eurostat/statistics-explained/index.php?oldid=497115#: :text=In%202022%2C%20the%20EU%27s%20working,by%202100%20(%2D9.5%20pp.)*

Fernández-Reino, M. and Kierans, D. (2020), 'Locking out the Keys? Migrant Key Workers and Post-Brexit Immigration Policies', *The Migration Observatory* **29**.

Fetzer, T., Rauh, C. and Schreiner, C. (2024), 'The Hidden Toll of the Pandemic: Excess Mortality in non-COVID-19 Hospital Patients', *Journal of Health Economics* **95**, 102882.

Fetzer, T. and Wang, S. (2020), 'Measuring the Regional Economic Cost of Brexit: Evidence Up to 2019', *CEPR Discussion Paper No. 15051* .

Financial Times (2019), 'Anxiety for Europeans Living in UK Grows as Brexit Date Approaches'. `https://www.ft.com/content/8835a90c-cff9-11e9-b018-ca4456540ea6`.

Foster, A. D. and Lee, Y. S. (2015), 'Staffing Subsidies and the Quality of Care in Nursing Homes', *Journal of Health Economics* **41**, 133–147.

Friedberg, R. M. (2001), 'The Impact of Mass Migration on the Israeli Labor Market', *Quarterly Journal of Economics* **116**(4), 1373–1408.

Friedberg, R. M. and Hunt, J. (1995), 'The Impact of Immigrants on Host Country Wages, Employment and Growth', *Journal of Economic Perspectives* **9**(2), 23–44.

Friedrich, B. U. and Hackmann, M. B. (2021), 'The Returns to Nursing: Evidence from a Parental-leave Program', *The Review of Economic Studies* **88**(5), 2308–2343.

Frost, D. M. (2020), 'Hostile and harmful: Structural Stigma and Minority Stress Explain Increased Anxiety Among Migrants Living in the United Kingdom After the Brexit Referendum.', *Journal of Consulting and Clinical Psychology* **88**(1), 75.

Furtado, D. and Ortega, F. (2023), 'Does Immigration Improve Quality of Care in Nursing Homes?', *Journal of Human Resources* . forthcoming.

Gallen, Y. (2019), 'The Effect of Parental Leave Extensions on Firms and Coworkers'. Working Paper.

Gaynor, M., Moreno-Serra, R. and Propper, C. (2013), 'Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service', *American Economic Journal: Economic Policy* **5**(4), 134–166.

Ginja, R., Karimi, A. and Xiao, P. (2023), 'Employer Responses to Family Leave Programs', *American Economic Journal: Applied Economics* **15**(1), 107–135.

Glover, D. and Henderson, J. (2010), 'Quantifying Health Impacts of Government Policies', *London: Department of Health* .

Grabowski, D. C., Gruber, J. and McGarry, B. (2023), Immigration, The Long-Term Care Workforce, and Elder Outcomes in the US, Technical report, National Bureau of Economic Research.

Gruber, J. and Kleiner, S. A. (2012), 'Do Strikes Kill? Evidence from New York State', *American Economic Journal: Economic Policy* **4**, 127–157.

Guriev, S. and Papaioannou, E. (2022), 'The Political Economy of Populism', *Journal of Economic Literature* **60**(3), 753–832.

Hantzsche, A., Kara, A. and Yong, G. (2019), 'The Economic Effects of the UK Government's Proposed Brexit Deal', *The World Economy* **42**(1), 5–20.

Hornung, E. (2014), 'Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia', *American Economic Review* **104**, 84–122.

House of Commons Library (2020), 'The EU Settlement Scheme: A Summary'. `https://commonslibrary.parliament.uk/the-eu-settlement-scheme-a-summary/`.

Huber, K., Lindenthal, V. and Waldinger, F. (2021), 'Discrimination, Managers, and Firm

Performance: Evidence from "Aryanization"', *Journal of Political Economy* **129**(9), 2455–2503.

Huebener, M., Jessen, J., Kuehnle, D. and Oberfichtner, M. (2022), 'A Firm-Side Perspective on Parental Leave'. Working Paper.

Jones, B. F. and Olken, B. A. (2005), 'Do Leaders Matter? National Leadership and Growth Since World War II', *Quarterly Journal of Economics* **120**(3), 835–864.

Kerr, W. R. and Lincoln, W. R. (2010), 'The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention', *Journal of Labor Economics* **28**(3), 473–508.

KPMG (2017), The Brexit Effect on EU Nationals, Technical report. `https://assets.kpmg.com/content/dam/kpmg/uk/pdf/2017/08/the-brexit-effect-on-eu-nationals.pdf`.

Kuhn, P. and Lizi, Y. (2021), 'How Costly is Turnover? Evidence from Retail', *Journal of Labor Economics* **39**(2), 461–469.

Lee, J., Peri, G. and Yasenov, V. (2022), 'The Labor Market Effects of Mexican Repatriations: Longitudinal Evidence from the 1930s', *Journal of Public Economics* **205**, 104558.

Manacorda, M., Manning, A. and Wadsworth, J. (2012), 'The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain', *Journal of the European Economic Association* **10**(1), 120–151.

Mitaritonna, C., Orefice, G. and Peri, G. (2017), 'Immigrants and Firms' Outcomes: Evidence from France', *European Economic Review* **96**, 62–82.

Moscelli, G., Gravelle, H. and Siciliani, L. (2021), 'Hospital Competition and Quality for Non-emergency Patients in the English NHS', *The RAND Journal of Economics* **52**(2), 382–414.

Moscelli, G., Gravelle, H., Siciliani, L. and Gutacker, N. (2018), 'The Effect of Hospital Ownership on Quality of Care: Evidence from England', *Journal of Economic Behavior & Organization* **153**, 322–344.

NHS Digital (2023*a*), NHS Workforce Statistics - January 2023, Technical report. Technical report. `https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/january-2023`. Accessed on 06-11-2023.

NHS Digital (2023*b*), Summary Hospital-level Mortality Indicator (SHMI), Technical report. Technical report. `https://digital.nhs.uk/data-and-information/publications/ci-hub/summary-hospital-level-mortality-indicator-shmi`. Accessed on 10-06-2023.

Nursing Times (2018), 'Lack of post-Brexit Immigration Plan Causing EU Nurses 'Anxiety''. `https://www.nursingtimes.net/news/workforce/lack-of-post-brexit-immigration-plan-causing-eu-nurses-anxiety-14-02-2018/`.

OECD (2017), *Health at a Glance 2017*, OECD Publishing.

OECD (2021), 'Pensions at a Glance'.
URL: *https://www.oecd-ilibrary.org/docserver/ca401ebd-en.pdf?expires=1690561872&id=id&accname=guest&checksum=32FCFD40C44E721A8721CD33F67DDFA3*

Ottaviani, G. I. P., Peri, G. and Wright, G. C. (2018), 'Immigration, Trade and Productivity in Services: Evidence from U.K. firms', *Journal of International Economics* **112**, 88–108.

Peri, G. (2012), 'The Effect of Immigration on Productivity: Evidence from US States', *Review of Economics and Statistics* **94**(1), 348–358.

Peri, G. (2014), 'The Labour Market Effects of Immigration and Emigration in OECD countries', *The Economic Journal* **124**(579), 1106–1145.

Peri, G., Shih, K. and Sparber, C. (2015), 'STEM Workers, H-1B Visas, and Productivity', *Journal of Labor Economics* **33**, S225–S255.

Portes, J. (2022), 'Immigration and the UK Economy after Brexit', *Oxford Review of Economic Policy* **38**(1), 82–96.

Propper, C. and van Reenen, J. (2010), 'Can Pay Regulation Kill? Panel Data Evidence on the Effect of Labor Markets on Hospital Performance', *Journal of Political Economy* **118**, 222–273.

Rodrik, D. (2021), 'Why Does Globalization Fuel Populism? Economics, Culture, and the Rise of Right-wing Populism', *Annual Review of Economics* **13**(1), 133–170.

Ryen, L. and Svensson, M. (2015), 'The Willingness to Pay for a Quality Adjusted Life Year: a Review of the Empirical Literature', *Health Economics* **24**(10), 1289–1301.

Sumption, M. and Fernandez Reino, M. (2018), 'Exploiting the Opportunity? Low-skilled Work Migration after Brexit', *Migration Observatory, University of Oxford* .

Teodorowski, P., Woods, R., Czarnecka, M. and Kennedy, C. (2021), 'Brexit, Acculturative Stress and Mental Health among EU Citizens in Scotland', *Population, Space and Place* **27**(6), e2436.

Terry, S. J., Chaney, T., Burchardi, K. B., Tarquinio, L. and Hassan, T. A. (2023), 'Immigration, Innovation, and Growth'. Working Paper.

The Guardian (2019), 'More EU Citizens are Seeking Help for Stress and Anxiety over Brexit'. `https://www.theguardian.com/politics/2019/jun/02/eu-citizens-seeking-help-stress-anxiety-brexit`.

UK Government (2020), 'The UK's Points-based Immigration System: Information for EU citizens'. `https://www.gov.uk/guidance/the-uks-points-based-immigration-system-information-for-eu-citizens`.

U.S. Bureau of Labor Statistics (2023), 'The Beveridge Curve (Job Openings Rate vs. Unemployment Rate), Seasonally Adjusted'.
URL: *https://www.bls.gov/charts/job-openings-and-labor-turnover/job-openings-unemployment-beveridge-curve.htm*

Waldinger, F. (2010), 'Quality Matters: The Expulsion of Professors and the Consequences for PhD Student Outcomes in Nazi Germany', *Journal of Political Economy* **118**(4), 787–831.

Waldinger, F. (2012), 'Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany', *Review of Economic Studies* **79**(2), 838–861.

Eidesstattliche Versicherung

Ich, Herr Kai Fischer, versichere an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf erstellt worden ist.

Duisburg, der 21. März 2025 Unterschift: ______________________________

Statements of Contribution

Statement of Contribution

My co-authors Simon Martin and Philipp Schmidt-Dengler contributed to the chapter

‘The Heterogeneous Effects of Entry on Prices’

of my dissertation

‘Five Essays on Industrial Organization and Economic Policy’.

All authors contributed *equally* to this chapter.

Signature of Simon Martin: ____

Statement of Contribution

My co-authors Simon Martin and Philipp Schmidt-Dengler contributed to the chapter

‘The Heterogeneous Effects of Entry on Prices’

of my dissertation

‘Five Essays on Industrial Organization and Economic Policy’.

All authors contributed *equally* to this chapter.

Signature of Philipp Schmidt-Dengler: ____

Statement of Contribution

My co-authors Simon Martin and Philipp Schmidt-Dengler contributed to the chapter

‘Indirect Taxation in Consumer Search Markets: The Case of Retail Fuel’

of my dissertation

‘Five Essays on Industrial Organization and Economic Policy’.

All authors contributed *equally* to this chapter.

Signature of Simon Martin: ____

Statement of Contribution

My co-authors Simon Martin and Philipp Schmidt-Dengler contributed to the chapter

'Indirect Taxation in Consumer Search Markets: The Case of Retail Fuel'

of my dissertation

'Five Essays on Industrial Organization and Economic Policy'.

All authors contributed *equally* to this chapter.

Signature of Philipp Schmidt-Dengler: ____

Statement of Contribution

My co-authors Henrique Castro-Pires, Marco Mello, and Giuseppe Moscelli contributed to the chapter

‘Immigration, Workforce Composition, and Organizational Performance:

The Effect of Brexit on NHS Hospital Quality’

of my dissertation

‘Five Essays on Industrial Organization and Economic Policy’.

All authors contributed *equally* to this chapter.

Signature of Henrique Castro-Pires: ____

Statement of Contribution

My co-authors Henrique Castro-Pires, Marco Mello, and Giuseppe Moscelli contributed to the chapter

> 'Immigration, Workforce Composition, and Organizational Performance:
> The Effect of Brexit on NHS Hospital Quality'

of my dissertation

> 'Five Essays on Industrial Organization and Economic Policy'.

All authors contributed *equally* to this chapter.

Signature of Marco Mello: ____

Statement of Contribution

My co-authors Henrique Castro-Pires, Marco Mello, and Giuseppe Moscelli contributed to the chapter

'Immigration, Workforce Composition, and Organizational Performance: The Effect of Brexit on NHS Hospital Quality'

of my dissertation

'Five Essays on Industrial Organization and Economic Policy'.

All authors contributed *equally* to this chapter.

Signature of Giuseppe Moscelli: ____