Empirical Analyses of Societal Challenges: Social Cohesion, Labor Market Transition and Population Health

Universität Hamburg
Faculty of Business, Economics and Social Sciences

DISSERTATION

to obtain a doctorate in
business, economics, and the social sciences
"Dr. rer. pol."
(Pursuant to the doctoral degree regulations of 18 January 2017)

submitted by

Shushanik Margaryan
from Artashat (Armenia)

Hamburg, 25 February 2021

Chairperson
Prof. Dr. Johanna Kokot, Universität Hamburg

First examiner:
Prof. Thomas Siedler (PhD), Universität Hamburg

Second examiner:
Prof. Dr. Moritz Drupp, Universität Hamburg

Third examiner:
Prof. Dr. Miriam Beblo, Universität Hamburg

The disputation was held on: 7 September 2021

To my grandmother Rima, in loving memory

Acknowledgments

This dissertation was written during my time as a research and teaching assistant at the Universität Hamburg. Several people enabled this journey. Above all, I am indebted to my supervisor Thomas Siedler for his unparalleled academic and personal support throughout these years. This dissertation would not have been possible without his sharp advice, gentle nudge towards self-betterment and encouragement. I am also thankful to him for providing a positive and wholesome environment at the chair, and the trust he showed during this process.

Throughout these years I have learned and benefited from multiple members of the economics department and the Hamburg Center for Health Economics. I would like to particularly thank Jan Marcus for his intellectual input and witty jokes, Miriam Beblo for her valuable and veracious feedback, Jonas Schreyögg for the welcoming atmosphere at HCHE.

An inseparable part of this journey have been many colleagues who shared the joy and the pain of writing a dissertation, provided a safe place to discuss, grow and distract. I would like to particularly thank Michael Bahrs and Raffael Kamalow for being excellent "chairmates" to whom I could turn with the mundane academic questions, Luise Görges and Eva Markowsky for the support, the laughs and for learning together how to navigate academia as female economists. I am also thankful to my fellow colleagues Pier, Luciana, Christos, Daniel, Annika, Stefan, Jakob and many others.

During my PhD, I was lucky to meet colleagues who not only positively influenced my research, but also became friends. I would like to particularly thank Alexandra de Gendre and Nicolas Salamanca for hosting me at the Melbourne Institute of Economic Research and for our countless enlightening discussions. I further want to thank Armando Meier and Alexander Ahammer for the valuable feedback they provided on numerous occasions.

My special thanks go to Arthur Kochnakyan for inspiring me to pursue economics, to my friends, Merri, Mattis, Katharina, Sören, and many others who have been paramount in sustaining my mental health during these times.

Lastly, I want to thank my family, and most importantly my grandmother, who was my source of endless and unconditional support.

Abstract

This dissertation explores three dimensions of societal well-being. Chapter 1 focuses on social cohesion and its relationship with schooling. Chapter 2 analyzes the effect of internships in terms of young people's transition to the labor market. Chapter 3 investigates population health and air pollution in the context of urban environmental cohesion. The three chapters of this dissertation are connected through the overarching topic of societal well-being and through the application of causal empirical methods. Each chapter is, nevertheless, self-contained and seeks to answer a specific research question.

Chapter 1 shows that an additional year of schooling reduces the probability of individuals and their children being very concerned about immigration. To isolate the exogenous variation in years of schooling, we use the staggered extension of compulsory schooling in West Germany between 1949 and 1969 in an instrumental variable design. The study estimates that an additional year of education reduces the probability of being concerned about immigration by around six percentage points. We also document large intergenerational spillovers. Furthermore, exploring an array of potential mechanisms, we find that labor market outcomes do not appear to be mediators of this effect. Instead, we show that social trust increases with an additional year of schooling, and this higher social trust potentially explains the lower level of concern about immigration.

Chapter 2 studies the effect of university internships on the labor market outcomes of young people up to five years after their graduation. This study exploits variation in the introduction and the discontinuation of mandatory internships within a field of study. We use two instrumental variable approaches. First, we

instrument internship completion with mandatory internships. As a second instrumental variable, we use the share of students reporting mandatory internships for their cohort and field of study in a leave-one-out fashion. Both instrumental variables imply large returns to internships. On average, our estimates suggest that students who completed an internship earn 6–14 percent higher wages after graduation. The findings also show that students who completed an internship face a lower risk of unemployment. The return of internships is especially pronounced for students who graduate from a field of studies with a weak labor market orientation.

Chapter 3 shows that targeted urban environmental policies may reduce air pollution and may have a positive impact on population health. Using the considerable variation in the timing of the implementation of low emission zones in German cities, this study shows that particulate matter (PM_{10}) and nitrogen dioxide (NO_2) concentrations decreased in treated cities. Further, analysis of the outpatient health diagnosis data shows that due to this decrease in pollution, the number of people diagnosed with cardiovascular disease declined by 2–3 percent. Slicing the population by age groups, I show that those over the age of 65 especially benefited from the policy.

German Summary (Zusammenfassung)

Diese Dissertation untersucht drei Dimensionen des gesellschaftlichen Wohlbefindens: Kapitel 1 konzentriert sich auf die Beziehung zwischen sozialem Zusammenhalt und Schulbildung. Kapitel 2 analysiert die Auswirkung von Berufspraktika im Rahmen des Einstiegs junger Menschen in den Arbeitsmarkt. Kapitel 3 untersucht den Zusammenhang zwischen städtischer Umweltpolitik, vor allem im Bereich Luftverschmutzung, und der örtlichen Bevölkerungsgesundheit. Die drei Kapitel dieser Dissertation sind durch das zugrundeliegende Thema des gesellschaftlichen Wohlbefindens und durch die Anwendung kausaler empirischer Methoden miteinander verbunden. Jedes Kapitel ist jedoch in sich abgeschlossen und folgt einer spezifischen Forschungsfrage.

Kapitel 1 zeigt, dass die Wahrscheinlichkeit einer Besorgnis über Zuwanderung bei Einzelpersonen und ihren Kindern reduziert wird, wenn ein zusätzliches Schuljahr absolviert wurde. Um die exogene Variation der Schuljahre zu isolieren, nutzen wir die zeitlich versetzte regionale Verlängerung der Schulpflicht in Westdeutschland zwischen 1940-1960 in einem Instrumentalvariablendesign. Die Studie zeigt, dass ein zusätzliches Schuljahr die Wahrscheinlichkeit, über Zuwanderung besorgt zu sein, um etwa sechs Prozentpunkte reduziert. Wir stellten zudem große intergenerationale Spillover-Effekte fest. Bei der Untersuchung einer Reihe potenzieller Mechanismen zeigte sich, dass die Arbeitsmarktergebnisse keine offensichtlichen Mediatoren für diesen Effekt sind. Stattdessen zeigen wir, dass das soziale Vertrauen mit einem zusätzlichen Schuljahr zunimmt. Dieses höhere soziale

Vertrauen erklärt möglicherweise die geringeren Bedenken im Hinblick auf Zuwanderung.

Kapitel 2 untersucht die Auswirkung von Berufspraktika während des Studiums auf die Arbeitsmarktergebnisse junger Menschen, und zwar bis zu fünf Jahre nach ihrem Studienabschluss. Diese Studie nutzt Variation hinsichtlich der Einführung und Abschaffung von Pflichtpraktika innerhalb eines Studienfachs. Wir verwenden zwei Instrumentalvariablenansätze. Zuerst instrumentieren wir die Absolvierung von Praktika mit Pflichtpraktika. Als zweites Instrument verwenden wir den Anteil der Studierenden, die Pflichtpraktika für ihre Kohorte und ihr Studienfach angeben, in einer Leave-One-Out-Methode. Beide Instrumentalvariablen implizieren große Renditen für Praktika. Im Durchschnitt deuten unsere Schätzungen darauf hin, dass Studierende, die ein Praktikum absolviert haben, nach ihrem Abschluss zwischen 6-14% höhere Löhne aufweisen. Die Ergebnisse zeigen auch, dass Studierende, die ein Praktikum absolviert haben, ein geringeres Risiko der Arbeitslosigkeit haben. Der Ertrag von Praktika ist besonders ausgeprägt für Studierende, die einen Studiengang mit schwacher Arbeitsmarktorientierung absolvieren.

Kapitel 3 zeigt, dass eine gezielte städtische Umweltpolitik die Luftverschmutzung reduzieren und einen positiven Einfluss auf die Gesundheit der Bevölkerung haben kann. Anhand der sehr unterschiedlichen Zeitpunkte der Einführung von Umweltzonen in deutschen Städten zeigt diese Studie, dass die Feinstaub- (PM_{10}) und Stickstoffdioxid-Konzentrationen (NO_2) in den Städten mit eingeführter Umweltzone abnahmen. Die Analyse der Daten zu den ambulanten Gesundheitsdiagnosen zeigt außerdem, dass aufgrund dieser Abnahme der Schadstoffbelastung die Zahl der Personen, bei denen eine Herz-Kreislauf-Erkrankung diagnostiziert wurde, um 2-3% zurückging. Unterteilt man die Bevölkerung nach Altersgruppen so zeigt sich, dass vor allem die über 65-Jährigen von der Maßnahme profitierten.

List of Publications

Margaryan, Shushanik, Paul, Annemarie, Siedler, Thomas. (2019). Does education affect attitudes towards immigration? Evidence from Germany. *Journal of Human Resources*, 0318-9372R1.

Margaryan, Shushanik, Saniter, Nils, Schumann, Mathias, Siedler, Thomas (2020). Do internships pay off? The effects of student internships on earnings. *Journal of Human Resources*, 0418-9460R2.

Margaryan, Shushanik (2021). Low emission zones and population health. *Journal of Health Economics*, 76, 102402.

Contents

General Introduction

> *Society does not find the bases on which it rests fully laid out in consciences; it puts them there itself.*
>
> Émile Durkheim

Motivation

Measuring the well-being of societies is a complex and intricate task. While gross domestic product (GDP) is traditionally used by economists and policymakers as the standard measure of a country's performance, there is a burgeoning debate on the caveats of this metric. GDP does not reflect factors such as environmental damage, growing inequalities, unfavorable labor market conditions for young people, or deteriorating trust in governments. The aftermath of the 2008 financial crisis ultimately demonstrated that measuring societal well-being using the simple metric of aggregate production can give rise to misleading conclusions and a spurious impression of prosperity, as the determinants of societal well-being are in fact multi-dimensional, dynamic, and adaptable. There have been several attempts to develop indicators to complement or replace GDP. Arguably the most renowned attempt is the report by the Commission on the Measurement of Economic Performance and Social Progress (Stiglitz-Sen-Fitoussi Commission), a group set up by France's former president Nicolas Sarkozy in 2008. The commission identifies eight key dimensions of well-being: material living standards, health, education, personal activities including work, political voice and governance, social connections and relationships, environment, and insecurity (Stiglitz et al., 2009). The Stiglitz-Sen-Fitoussi report sparked a vibrant political and academic debate, inspired the Organisation for Economic Co-operation and Development (OECD) to develop its "Better Life Index", and prompted the European Statistical System to translate its recommendations into concrete indicators to measure progress.

The Stiglitz-Sen-Fitoussi report also influenced the EU's development agenda in the aftermath of the financial crisis. In its 2009 report "GDP and beyond: Measuring progress in a changing world", the European Commission emphasizes the need to complement GDP with environmental and social indicators. The EU's long-term development plan for the period 2010-2020, the "Europe 2020" strategy, noticeably departs from the approach of using GDP as the sole policy goal. Instead, the strategy concentrates on thematic objectives with a view to ensuring "smart, sustainable, and inclusive growth".

One instrument employed to help achieve the goals of "Europe 2020" is "Horizon 2020", a framework program funding research in key areas. Within the program, the EU defines seven priority areas, which it calls societal challenges. This dissertation focuses on the first block of these challenges, namely "Health, demographic change and wellbeing". Each chapter concentrates on a specific dimension of this priority area. That said, the areas being considered are all interconnected, integral elements of societal well-being as a whole. Chapter 1 explores the relationship between schooling and social cohesion, measured through political attitudes. Chapter 2 addresses the topic of higher education and young people's transition into the labor market. Chapter 3 focuses on air pollution and health outcomes.

The first chapter of this dissertation examines two key dimensions of social cohesion: attitudes toward immigration and generalized trust. Social cohesion has received significant attention from policymakers over the last two decades. The EU and the Council of Europe made it a strategic priority as early as 2000. However, despite the significant academic and political interest in social cohesion, there is no consensus on its definition. In their review of the literature, Schiefer and van der

Noll (2017) identify six essential dimensions of cohesion across numerous definitions. The authors show that social relations between individuals and between various groups form the foundation of social cohesion. If these groups differ, such relations require mutual tolerance, making tolerance an integral component of social cohesion. Furthermore, as Émile Durkeim already noted in 1893, the transformation of societies changes the nature of cohesion (Durkheim, 2014). The constant demographic and political change in many EU Member States made xenophobia and intolerance toward immigrants a central threat to social cohesion in the union. In 2016, the EU launched the Action Plan "Building Inclusive Societies" centered around three areas: education, combating intolerance, and integration policies.

After decreasing in relevance in the aftermath of World War II, from the 2000s, right-wing populist parties flooded the political scene of almost all major economies. While the economic rhetoric of these parties differs from country to country, their political rhetoric universally singles out immigrants as a scapegoat, a significant cause of various economic, social and political challenges a country faces. The rise of populism was exacerbated in Europe, and especially in Germany, when in 2015, around 1.3 million refugees sought asylum in different EU Member States. Germany's chancellor Angela Merkel famously opened Germany's borders to nearly one million refugees, a political move that lost her a significant number of supporters, as well as prompted severe criticism and heated debates in the country. Accompanying all these events was the plummeting trust in the establishment, captured by British politician Michael Gove's public announcement that the "people have had enough of experts".

The anti-migration rhetoric of right-wing parties helped to summon votes for the "Leave" vote in the UK's Brexit referendum and to elect Donald J. Trump, with his harsh anti-migration position, as the 45th president of the US. In the EU, multiple right-wing parties obtained seats in parliament in countries where they had previously been unsuccessful. In Italy, the anti-establishment Five Star Movement came to power in a coalition with the far-right, anti-immigration League party, while populist Marine Le Pen had most of western Europe holding its breath in her runoff against Emanuel Macron in 2017. Germany, a country with a tragic history of radical right-wing politics, was shaken by the increasing popularity of the far right Alternative für Deutschland (AfD), which gained 92 seats in the German Bundestag in 2017.

What determines individual's attitudes towards immigration? Decades of economic and sociological literature find correlations with an array of factors, including skills, employment and income (Scheve and Slaughter, 2001; Mayda, 2006), education (Coenders and Scheepers, 2003; O'Rourke and Sinnott, 2006; Hainmueller and Hopkins, 2014; Lancee and Sarrasin, 2015) welfare concerns and prejudice (Dustmann and Preston, 2001), neighborhood conditions (Van Heerden and Ruedin, 2019), and life satisfaction (Poutvaara and Steinhardt, 2018). From a policy perspective, however, the important question is what causes these attitudes and what can lead them to change. With a view to exploring one of the factors in this context, Chapter 1 sets out to investigate the causal effect of education on attitudes toward immigration.

While a cohesive society undoubtedly requires stable foundations in terms of civic education, ensuring cohesion requires economic opportunities and functioning labor markets, as well. Hence, the second chapter of this dissertation considers young people's transition to the labor market. The consistently high youth unemployment rate is one of the most persistent problems the EU has faced over the last decade, a problem that has been exacerbated by the financial crisis. While not all Member States have equally high levels of youth unemployment, the persistent gap between the general and the youth unemployment rate is a common problem. The average difference in the 28 EU Member States was as high as 13 percentage points in 2011. And while, compared to other European countries, Germany has one of the lowest youth unemployment rates, the gap between the general unemployment rate and the youth unemployment rate has still been as much as around three percentage points since 2011 (Eurostat, 2020).

There are multifaceted macroeconomic and microeconomic reasons behind this phenomenon. Alongside country-specific factors, some issues are universal. One such problem is the arduous transition from education to the labor market, which was made even more difficult by the European financial crisis. Since the onset of the crisis, even the young people with the highest level of education face increasingly difficult transitions (Eurofound, 2012).

The transition experience can have a long-lasting impact on the career trajectories of young people. Fredriksson et al. (2018), for instance, show that an initial job mismatch can have significant consequences on future labor market outcomes and that the magnitude of this impact depends on the information available before hiring. Ruhm (1991) and Arulampalam (2001) show that unemployment can have scarring effects on future labor market outcomes. Moreover, young people who graduate from areas of study with a weak labor market orientation, such as humanities, face a more difficult transition and subsequently lower remuneration (Arcidiacono, 2004; Grave and Görlitz, 2012).

Acquiring the necessary experience and skills is key to facilitating this transition, and education and training systems play a vital role in making this possible. Governments around the world have initiated efforts at different levels of education to create a closer link between education and the labor market. In 2013, the EU rolled out the Youth Guarantee scheme, which obliges the Member States to present implementation plans to help young people find employment. Internships are major part of these implementation plans. In 2009, the UK launched the Graduate Talent Pool, a service that provides recent graduates with a chance to participate in a paid internship. The START program provides a similar service in Romania. In France, education establishments are formally assigned the role of organizing internships.

Following the changes to education implemented as part of the Bologna Reform, graduates' employability has become a key objective of higher education across Europe (Teichler, 2011). Universities were called upon to increase the focus on providing students with labor market-relevant skills and preparing graduates for the transition to work. Internships are an effective means of developing these skills (Wolter and Banscherus, 2012; Teichler, 2011). As a result, the rates of internship take-up have soared. In Germany, by the time students finish their studies, nearly 80 percent report to have completed an internship while at university (Sarcletti, 2009). But do these internships pay off? This is the question Chapter 2 aims to answer through analyzing the wages and the employment probability of young people who completed an internship during their studies.

The final chapter of this dissertation zeroes in on the challenge of ensuring a healthy urban environment. In 2007, for the first time, the number of people living in urban areas surpassed the number living in rural areas (UN, 2018). In 2017, 75% of the EU's population lived in urban areas, a number that has been on the rise ever since. This rapid urbanization, alongside the manifold opportunities it creates, also poses multiple problems for cities, however. One central issue is the increasing number of road vehicles: Over the last five years, the EU passenger car fleet grew by eight percent. In 2019, there were, on average, 569 cars per 1,000 inhabitants in the EU (European Automobile Manufacturers Association, 2021). Road traffic is one of the largest emitters of air pollution and greenhouse gases in urban areas. It is the largest source of nitrogen oxides (NO and NO_2) emissions, and the second largest source of carbon monoxide (CO) and particulate matter (PM) emissions. Urban areas also account for 23% of CO_2 emissions from transport (European Environment Agency, 2016).

The high congestion and pollution levels caused by the large number of road vehicles are a challenge from the perspective of ensuring healthy environments and socio-environmental equality. A significant share of the urban population in the EU is exposed to pollution levels that exceed EU air quality thresholds. The European Commission estimates that the total health-related costs of air pollution amount to several hundred billion euros per year (Science for Environment Policy, 2021).

Medical and epidemiological studies have long shown the detrimental impact of pollution on both physical and mental health (Linn et al., 2000; Pope, 2000; Tsai et al., 2003; Brook et al., 2004; Wellenius et al., 2005; Franchini and Mannucci, 2007; Fiordelisi et al., 2017; Xue et al., 2019; Bolton et al., 2013; Buoli et al., 2018). The economics literature contributes to the evidence on the effects of pollution by focusing on causality and explicitly recognizing how sorting behaviors in the form of residential preferences and avoidance behavior can exacerbate existing health inequalities and bias the estimates (Graff Zivin and Neidell, 2013).

Despite this extensive literature, the evidence on what works in terms of policy is scarce. Davis (2008) shows that license plate-based driving restrictions in Mexico have not improved air quality, a finding that has been replicated in Santiago (Gallego et al., 2013) and Bogota (Zhang et al., 2017), while the findings from a similar policy intervention for Beijing are contested (Viard and Fu, 2015; Sun et al., 2014). Auffhammer and Kellogg (2011) find that gasoline content regulations had no effect in the US. Other forms of regulations such as scrappage subsidies, registration fees and mandatory catalytic converters have delivered mixed results (Greenstone and Hanna, 2014). The evidence is more encouraging when it comes to congestion charges: Green et al. (2020) find substantial positive effects in London, Gibson and Carnovale (2015) in Milan, and Simeonova et al. (2018) in Stockholm.

In Europe, low emission zones are the most frequently implemented instruments to address vehicle-generated pollution. Currently, there are around 250 LEZs across EU Member States. Chapter 3 sets out to evaluate the effect of this measure on air pollution and population health, exploiting the favorable institutional setting in Germany.

Overview of Thesis Chapters and Contributions

As mentioned in the previous subsection, this thesis includes three main chapters. While they all seek to address the same overarching research question and are connected methodologically, each chapter investigates a distinct sub-field and is self-contained. In the following pages, I present a succinct discussion of each chapter and its primary contributions to the economics literature.

Chapter 1 of the thesis is a joint study with Annemarie Paul and Thomas Siedler. This chapter investigates the causal relationship between education and immigration attitudes in Germany. We exploit an extension of compulsory schooling implemented in West German federal states between 1949 and 1969. The study finds that additional schooling reduces the probability of individuals reporting being very concerned about immigration. The paper makes several important contributions. First, this is the first study to comprehensively examine the effect of schooling on an individual's own attitudes towards immigration and those of their offspring. The intergenerational aspect is particularly important in quantifying the composite non-monetary returns to education. The study shows that considering only individual returns understates the total impact of such educational policies. Second, this study provides novel evidence on the mechanisms that mediate the effect of schooling. We start by confirming the findings from the previous literature that labor market outcomes are unlikely to explain the relationship between schooling and attitudes. Next, we demonstrate that social capital may play a substantial role in explaining our findings by showing the positive and significant effect of schooling on generalized trust. Third, the economics literature has traditionally focused on the economic returns to education. With Milligan et al. (2004) and Dee (2004), a new strand of literature emerges, focusing on the civic outcomes associated with education. Chapter 1 contributes to the relatively small body of literature on civic returns to education in general, and on the impact of schooling on attitudes toward immigration in particular. Finally, we also consult publications and historical documents to shed light on the curriculum of the ninth mandatory year of schooling. We document that the additional curriculum had two areas of focus: civic education and occupational education.

Chapter 2 is another joint study conducted in collaboration with Nils Saniter, Mathias Schumann, and Thomas Siedler. In this study, we investigate whether completing an internship while studying has an effect on the labor market outcomes of the student. Using mandatory internships at German universities as an instrument, we show that internships have large positive effects on the wages and employment probability of students five to six years after graduation. A noteworthy finding of this paper is that the internships disproportionately benefit the students from areas of study that have a weak labor market orientation. The study contributes to the economics literature by providing the first evidence on the causal effect of internships on the labor market outcomes of students, and also complements the large body of empirical literature on the returns to education for individuals with a high level of education.

Chapter 3 seeks to examine an urban environmental policy targeting air pollution in large cities in Germany. The study shows that the low emission zones introduced by the cities reduce the air pollution inside the zones and, through this reduction, also benefit the cardiovascular outcomes of the adult population, particularly those over the age of 65. This paper contributes to the existing literature in several ways. First, it adds to the small strand of literature studying the health effects of policy instruments targeting air pollution. Second, whereas most of the existing literature focuses on the respiratory health of infants and children, this paper enhances the literature by providing evidence for cardiovascular disease for the entire population and for the elderly. Studying the implications of such policies for the elderly is of particular interest because of their susceptibility to and the prevalence of circulatory system diseases. Third, this study complements the small number of existing studies by analyzing the effects of marginal long-term reductions in pollution in a setting with relatively low pollution levels. By exploiting the long-term variation, the findings in this paper complement the previous findings based on short-term temporal variations in pollution.

The final chapter of this thesis presents concluding remarks and policy implications.

Chapter 1

Does Education Affect Attitudes towards Immigration? Evidence from Germany[1]

Abstract

Using data from the German Socio-Economic Panel and exploiting the staggered implementation of a compulsory schooling reform in West Germany, this article finds that an additional year of schooling lowers the probability of being very concerned about immigration to Germany by around six percentage points. Furthermore, our findings imply significant spillovers from maternal education to immigration attitudes of her offspring. While we find no evidence for returns to education within a range of labour market outcomes, higher social trust appears to be an important mechanism behind our findings.

[1]This chapter is joint work with Annemarie Paul and Thomas Sidler. I gratefully acknowledge the helpful discussions with and comments of Silke Anger, Michael Bahrs, Miriam Beblo, Lorenzo Capellari, Wolfgang Dauth, Christian Dustmann, Colin Green, Jan Marcus, Regina Riphahn, Mathias Schumann, Ulf Zölitz and the participants of the 3rd BIEN annual conference, the 7th Ifo and TU Dresden Workshop on Labour Economics and Social Policy and the 9th IAB workshop on Perspectives on (Un-)Employment. This work was supported by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) through the project *"Nicht-monetäre Erträge von Bildung in den Bereichen Gesundheit, nicht-kognitive Fähigkeiten sowie gesellschaftliche und politische Partizipation"*. Accepted for publication in the *Journal of Human Resources.*

1.1 Introduction

Public concerns over immigration are increasingly used by populist political forces to mobilize voters. Framing immigrants as a threat to national-cultural identity, to the welfare system and to employment is an important element of the populist-right's political agenda: the European Union membership referendum in the United Kingdom was dominated by immigration issues and the message "take back control of borders" (Goodwin and Heath, 2016, p. 328); US president Donald Trump based his campaign largely on promises to restrict immigration; the populist party Alternative für Deutschland (AfD) explains in its programme why Islam does not belong in Germany (Alternative für Deutschland, 2016). Moreover, the strong prevalence of anti-immigration politics in the programs of far right-wing parties has led some scholars to view anti-immigration and racism as the main reasons why these parties have established themselves in many Western European countries (Rydgren and Ruth, 2011).

The increasing spread of populist political ideas raises the question whether education is an important factor that explains the demarcation of societies by their attitudes towards immigration. A positive correlation between education and pro-immigration attitudes is observed in voting polls, discussed excessively in the media, and documented in the scientific research (Scheve and Slaughter, 2001; Mayda, 2006; Hainmueller and Hiscox, 2007; Card et al., 2012). However, a causal relationship between education and pro-immigration attitudes is difficult to establish because of potential omitted variables. For instance, if individuals with highly educated parents have higher levels of education, and also inherit positive parental attitudes towards immigration, the OLS estimate of education is likely to be upwardly biased.

To estimate the effect of schooling on attitudes towards immigration, the present study uses the staggered implementation of a compulsory schooling reform in West Germany as a source of exogenous variation.

Our instrumental variable (IV) estimates indicate that an additional year of schooling reduces the probability of having high concerns about immigration to Germany by around six percentage points (20 percent). We further analyze potential intergenerational effects of maternal education on the offspring's immigration attitudes. The findings suggest that returns to education are not limited to the person directly affected, but also extend to the next generation. The probability of adult children being very concerned about immigration decreases by almost seven percentage points with an extra year of maternal schooling. This suggests an important composite effect of both educational and attitudinal spillovers.

Our analysis suggests that labor market outcomes are not the central transmission channel, since we confirm the previous findings that an additional year of schooling does not generate significant positive effects on outcomes such as labor force participation and labor income in Germany (Pischke and von Wachter, 2008; Kamhöfer and Schmitz, 2016). The analysis further reveals a positive impact of additional schooling on social capital, measured by general trust. We discuss the latter as a significant transmission channel. Due to data limitations, our ability to examine potential transmission channels of maternal schooling on offspring's attitudes is restricted. Nevertheless, in the light of the previous literature we argue that increased educational attainment of the offspring due to a higher level of maternal schooling and intergenerational correlations in immigration attitudes and trust are potentially relevant mechanisms.

The paper makes three main contributions to the literature. First, we examine the effect of schooling on individual's own *and* her offspring's attitudes towards immigration. The analysis of intergenerational effects, which to the best of our knowledge has not been done in other studies to date, is of particular interest because it has potentially important implications for the non-monetary returns to education: in presence of a positive effect of parents' education on children's attitudes, the net non-monetary returns are greater than individual estimates would suggest. As such, we are able to account for a more complete picture in terms of individuals' tolerance than previous studies. Second, given that the existing literature on mechanisms of education is mixed (Scheve and Slaughter, 2001; Dustmann and Preston, 2007; Hainmueller and Hiscox, 2007), we provide further evidence on relevant mechanisms, in particular labor market outcomes and the social capital of the individual. Third, we add to the relatively small literature on civic returns to education in general, and on the impact of schooling on attitudes towards immigration in particular.[2]

The remainder of the paper is organized as follows: Section 1.2 reviews the related literature, Section 1.3 provides the institutional background for Germany, and Section 1.4 presents the data. In Section 1.5, we discuss the empirical strategy and we present the results in Section 1.6. Section 1.7 discusses the potential transmission channels, and Section 1.8 concludes.

1.2 Related Literature

A large and growing body of literature examines the determinants of attitudes towards immigration. The studies most closely related to the present paper are those that explore the relationship between education and preferences over immigration. While the literature unequivocally agrees on a positive association between education and pro-immigration attitudes, the interpretation of factors behind this association differ.

One strand of studies attributes the association between education and immigration attitudes mainly to economic factors. Scheve and Slaughter (2001) find that less skilled individuals are more likely to prefer to restrict immigration. They use data from the American National Election Studies (NES) in the US for the

[2] While there exists a large literature on the monetary returns to education (see, for example, Angrist and Krueger, 1991; Card, 1999; Pischke and von Wachter, 2008; Kamhöfer and Schmitz, 2016; Oreopoulos, 2006), there is considerably less and mixed evidence on the civic returns to education (Dee, 2004; Milligan et al., 2004; Siedler, 2010; Lochner, 2011; Oreopoulos and Salvanes, 2011; Pelkonen, 2012). Furthermore, we are only aware of a single study (d'Hombres and Nunziata, 2016) that investigates the causal impact of education on attitudes towards immigration, without studying intergenerational effects however.

years 1992, 1994 and 1996 and measure skills both by educational attainment and average wage in the individual's occupation. To determine whether labour market competition between natives and immigrants explains their findings, the authors split the estimation sample between labour force participants and non-participants. If education predominantly measures labour market skills and immigrants are more likely to be low-skilled, the negative correlation between education and anti-immigration attitudes should only hold among labor force participants. By and large, this is what Scheve and Slaughter (2001) find. These results are corroborated by cross-country studies by Mayda (2006) and O'Rourke and Sinnott (2006) using data from the International Social Survey Programme (ISSP). Both studies reveal that more educated individuals are less likely to have anti-immigrant sentiments, and the association is stronger in richer countries where natives are more skilled than the immigrants.

Another strand of studies attributes the positive association between education and pro-immigration attitudes to non-economic factors. Arguably, education is related to other factors, such as openness and beliefs, that correlate with immigration attitudes (Hainmueller and Hopkins, 2014). Educational systems may convey tolerant and egalitarian values (Lancee and Sarrasin, 2015), enhance cognitive and analytical thinking to circumvent oversimplifications of social reality (Coenders and Scheepers, 2003), or even reduce the cost of analyzing political allegories to be less susceptible to hate-creating stories (Glaeser, 2005; Mocan and Raschke, 2016).

Related to this, Hainmueller and Hiscox (2007) argue that the conventional belief about labor market competition between natives and immigrants and anti-immigration sentiments is based on a misreading of the available evidence. The authors use data from the European Social Survey (ESS) with questions on attitudes towards immigration in four groups of country of origin: poor European, rich European, poor non-European and rich non-European. The reasoning is if labor market competition between natives and immigrants by skill level is the critical determinant of immigration preferences, education should be strongly and positively linked to support for immigration from poorer countries and more weakly, perhaps even negatively, correlated with the support for immigration from richer countries. The authors find that the highly educated are more likely to favor immigration irrespective of the immigrants' country of origin. Note, however, that the argument assumes immigrants to be a random sample from the source country, while there is strong evidence that immigrants are a self-selected group along a range of characteristics, including skills (Borjas, 1987). In contrast to Scheve and Slaughter (2001), Hainmueller and Hiscox (2007) do not find a significant difference in the association between education and attitudes towards immigration between the sub-samples of labor force participants and non-participants. Instead, they report that the more educated are less racist, place greater value on cultural diversity, and are more likely to believe that immigration brings benefits for the host economy as a whole.

Dustmann and Preston (2007) argue that across different levels of education, the aversion towards immigration is driven by various motives. The authors focus on three channels: welfare concerns, labor market concerns, and racial and cultural prejudices. Their findings suggest that welfare considerations are the most relevant factor in attitude formation of anti-immigration sentiments for the more educated, while racial and cultural prejudices play a dominant role for the less educated. At odds with the common expectation, the authors also find that for the less educated respondents labour market concerns play a minor role. This last finding should be interpreted and generalized with caution, since, unlike in the US and many European countries, UK immigrants, on average, have more schooling than the native-born whites (Dustmann and Preston, 2001).

Although extensive empirical research documents the positive association between education and pro-immigration attitudes, selection issues of individuals into higher education remain mostly unresolved. This point is raised by Lancee and Sarrasin (2015) in their recent analysis. The authors estimate the effect of education on attitudes towards immigration as adolescents pass through education in Switzerland, using a hybrid model that estimates a within-individual effect and a between-individual effect. The results confirm that education is the strongest between-individual predictor of attitudes towards immigration, but, when only within-person variance is used, the effect of education becomes statistically insignificant. This suggests that at least part of the positive relationship between education and pro-immigration attitudes is driven by selection into education.

The present study uses a compulsory schooling reform in West Germany to examine the effect of an additional year of schooling on immigration attitudes of native Germans.[3] In addressing the endogeneity issue of education, the present study is most closely related to recent work by d'Hombres and Nunziata (2016). The authors use data from the ESS and focus on a pool of 12 European countries.[4] For the whole sample of countries, they find a positive effect of education on pro-immigration attitudes in the order of 6-11 percentage points, on average. Considering that European countries differ greatly with respect to their institutional setting and historical development, the average effect for the joint sample of countries might hide important cross-country heterogeneity. Yet, single-country analysis remains an open dimension. Additionally, the authors use reforms across Europe, where schooling years differ in the base level of schooling to which the reform applies and the duration of additional schooling. For example, while schooling

[3] Starting with the seminal work by Angrist and Krueger (1991), there is a comprehensive body of literature that uses compulsory schooling laws as a source of exogenous variation. Lochner (2011) and Oreopoulos and Salvanes (2011) review the literature on non-pecuniary returns to education, and Holmlund et al. (2011) survey the intergenerational literature on the effects of parent's schooling on children's schooling. For a survey on pecuniary returns to schooling, see Card (1999). In the German context, studies analyze the effect of compulsory schooling on labor market outcomes (Pischke and von Wachter, 2008; Kamhöfer and Schmitz, 2016), political behavior (Siedler, 2010), health (Kemptner et al., 2011), fertility (Cygan-Rehm and Maeder, 2013), and the intergenerational transmission of education (Piopiunik, 2014). Moreover, exploiting exogenous variation in years of schooling allows us to address potential issues of reverse causality resulting from effects of immigration on the schooling of natives (McHenry, 2015; Hunt, 2017).

[4] d'Hombres and Nunziata (2016) include the following countries: Belgium, Denmark, Germany, Finland, France, Greece, Italy, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom (UK).

was increased from eight to 12 years in Belgium, it was increased from four to seven years in Italy. This asymmetry in the schooling reforms across countries generates distinct country-specific complier sub-populations for which an averaged local average treatment effect (LATE) is identified.

1.3 Institutional Background

1.3.1 Immigration to Germany

Germany is the largest migrant host country in Europe. By 2015, 21 percent of the total population had a migration background. The three most frequently represented migrant nationalities are Turkish, Polish and Italian (Bundesamt für Migration und Flüchtlinge, 2016).

The significant influx of foreign-born population began after the World War II and has been continuous ever since. An important group of immigrants after the war were ethnic Germans. According to the Basic Constitutional Law of the Federal Republic of Germany, this group can obtain German citizenship based on their ethnicity (Fassmann and Munz, 1994). Around 3.9 million ethnic Germans immigrated to Germany between 1950 and 1998.

In the 1960s, driven by the labor shortage, Germany initiated multiple bilateral agreements with southern European countries, primarily with Turkey and Italy. Worbs (2003) mentions that the number of Turkish guest workers in Germany rose from 2,495 in 1960 to 617,531 in 1974. By the end of 1973, migration policies in Germany were restrained due to oil price shocks and fears of economic recession. Nevertheless, the influx of immigrants continued at moderate rates mainly as a result of family reunification.

In the early 1990s, the influx of foreigners to Germany peaked due to a large number of asylum seekers and refugees, most of them fleeing from the war in former Yugoslavia. From 1989 to 1996, the annual total influx of foreign-born immigrants to Germany was consistently over one million, reaching roughly 1.6 million in 1992.

At present, Germany remains a major host country for economic migrants and refugees. A large proportion of economic migration is from other EU countries, mainly from recent member states, and as a result of the recession in the southern states. In 2015, the total number of immigrants reached 2.14 million, representing the largest number of immigrants since the start of the statistical recordings in 1950 (Bundesamt für Migration und Flüchtlinge, 2016). EU citizens account for 40 percent of the new arrivals, while by single country origin, Syrians and Romanians represent the largest group.

For further information about immigration to Germany since World War II see, for example, Riphahn (2003); D'Amuri et al. (2010); Glitz (2012); Braun and Kvasnicka (2014) and Braun and Mahmoud (2014).

1.3.2 The School System in Germany

Germany has a tripartite school system regulated at the federal state level. Nevertheless, the main features of the education system are almost identical across federal states (Dustmann, 2004). Compulsory school attendance begins at the age of six. Students spend the first four years at the primary school, after which they are tracked into three types of secondary schools: lower (*Hauptschule*, formerly called *Volksschule*), intermediate (*Realschule*) and academic (*Gymnasium*).[5] Currently, the *Hauptschule* requires attendance of nine years. It is the least demanding track and provides general education as a basis for apprenticeship training. The *Realschule* ends after the tenth class. This intermediate school track is more academically integrated and prepares students for both blue and white-collar jobs (Dustmann, 2004). The academic track—the *Gymnasium*—leads to a technical college entrance diploma (*Fachhochschulreife*) or to a university entrance diploma (*Abitur*). Obtaining an *Abitur* requires 12 to 13 years of schooling in all states in western Germany.[6]

[5] In Berlin and Brandenburg, the tracking takes place after the sixth class.

[6] Before 2001, students obtained their *Abitur* after 13 years at school. Since 2001, federal states introduced a school reform to shorten the length of high school by one year, resulting in

Historically, governance of school education was under state authority in Weimar Republic, until the Nazi Regime declared it to be under central organization in the beginning of 1934. As a result syllabus, teaching material and other school regulations fell largely under national control (Nicholls, 1978; Tent, 1982). With the emphasis on education and extracurricular activities loyal to the regime, the government promoted standardization of the school system and at least eight years of schooling. After World War II and the division of Germany into four occupation zones, federal state authorities regained decision power over schooling. The main focus was on re-establishing a continuous teaching and the denazification of educational content.

This study exploits a schooling reform in West Germany that extended compulsory schooling from eight to nine years. It was implemented over a span of 20 years—from 1949 to 1969—in different federal states at different points of time. This generates exogenous variation in years of schooling across states and over cohorts of students. Table 1.1 shows each federal state with the relevant reform date and year of birth of the first cohort to be affected by the education reform, taken from Pischke and von Wachter (2005).

In 1949, the first federal state, Hamburg, prolonged mandatory schooling from eight to nine years. Schleswig-Holstein, Bremen, Lower Saxony and Saarland followed within a time frame of fifteen years. Several debates and initiatives to further reform the school system accompanied. In 1964, states' striving to uniformly reform the school system led to passing the *Hamburg Accord*. Key agreements included the introduction of a nationwide nine-year compulsory education and the realization of two short school years under an unchanged curriculum. As late reformers Baden-Wuerttemberg, Hesse, North Rhine-Westphalia, and Rhineland-Palatinate extended compulsory schooling by one year in 1967 and, at the same time, postponed the beginning of the school year from the time around Easter

overall schooling of 12 years instead of 13 (Huebener and Marcus, 2017). However, this reform did not affect compulsory schooling.

to fall for two consecutive years.[7] Bavaria was the last state to introduce longer compulsory schooling in 1969.

1.4 Data and Descriptive Statistics

We use data from the Socio-Economic Panel study (SOEP v30)—a large representative longitudinal household survey that has interviewed around 11,000 households and 30,000 individuals annually since 1984 (Wagner et al., 2007). The SOEP contains a comprehensive set of socio-economic and demographic background variables as well as measures on personality traits, personal attitudes, and values.

We draw two samples. We use *Sample I* to study the direct effect of education on individuals' attitudes towards immigration. *Sample II* is drawn such that it allows us to analyze potential spillover effects of maternal education on their adult children's concerns about immigration.

In *Sample I*, we follow Pischke and von Wachter (2008) and Siedler (2010) and restrict our first estimation sample to cohorts born in Germany between 1930 and 1960, excluding individuals who either currently reside or attended school in the area of the former German Democratic Republic. Because the information on the federal state of school attendance is not available for all individuals in our sample, we use the first federal state of residence that appears in the SOEP as a proxy for state of school attendance if this information is not available.[8] Following Pischke and von Wachter (2008), we use the school leaving qualification, the information on year of birth and federal state to construct years of schooling as follows: those who have completed the academic track are assigned 12 or 13 years of schooling depending on whether they have a *Fachhochschulreife* or an *Abitur*, respectively. Those who followed the intermediate track are assigned ten

[7] Moreover, three other states, namely Schleswig-Holstein, Bremen, and Saarland, temporarily shortened the school years. We control for state-specific cohort trends in the main analysis and further discuss the influence of these short school years on our estimates in robustness section 1.6.3.

[8] In the robustness section, we relax this assumption and find similar results.

years of schooling. Those who completed the basic track and individuals with no school leaving qualification are assigned eight years of schooling if they were not exposed to the compulsory school reform and nine if they were. Accordingly, the compulsory school reform does not change the overall number of years of schooling for those who followed the intermediate or academic school track, while former students from the lowest school track have one additional year of schooling, depending on their year of birth and the state where they attended school.

Our main outcome describes individuals' attitudes towards immigration. We proxy this variable using the following question in the SOEP, which has been included every year since 1999: "Are you concerned about immigration to Germany?". The response categories are "very concerned", "somewhat concerned" and "not concerned at all". This measure of attitudes towards immigration has been previously used by Poutvaara and Steinhardt (2018) and Avdeenko and Siedler (2017).[9] We recode the original three-category variable into a dichotomous variable that equals one for individuals who report that they are "very concerned". Our outcome takes the value of zero for individuals who report that they are "somewhat concerned" or "not concerned at all".

The validity of our dependent variable depends on whether it in fact captures salience related to concerns people have about immigration. There are two good reasons for this to be the case. First, the question is clearly formulated as a concern. Lancee and Pardos-Prado (2013) argue that survey items that intend to measure public concern, capture the importance of the issue and the degree to which it is a problem, meaning that the respondents for whom concerns about immigration are salient are the ones who perceive it as a problem, or are unsatisfied with the current policy on the issue. Second, Avdeenko and Siedler (2017) compare the proportion of people who report being very concerned about immigration with the official share of votes for far right-wing parties at the federal state level in Germany.

[9]Note that Avdeenko and Siedler (2017) study intergenerational correlations in attitudes towards immigration, whereas the present paper aims at estimating the causal intergenerational spillovers of education on attitudes towards immigration.

They find a positive correlation coefficient of 0.40 between subjective measures on attitudes towards immigration as reported in the SOEP and the number of far right-wing votes in general elections.[10] Consequently, concerns about immigration are likely to translate into voting behavior that reflects individual intentions to reduce the concern.

In our main analysis, we use the first available observation for each individual. In the robustness section, we present an alternative specification exploiting the panel structure of the data: as an additional outcome, we use the proportion of years respondents are very concerned about immigration.

Sample II relies on retrospective information reported by adult children on the school leaving qualification and year of birth of their mother. We focus on mothers since the information reported about fathers can both refer to the biological father or to the mother's partner. Consistent with *Sample I*, *Sample II* comprises mothers born between 1930 and 1960 with valid information on their level of schooling.[11] Mothers' years of schooling and exposure to the reform are defined as in *Sample I*. The outcome captures adult children's concerns about immigration. The variable is dichotomous and is equal to one for adult children who report that they are "very concerned" about immigration and zero for children who report that they are "somewhat concerned" or "not concerned at all". In line with *Sample I*, we use the first available observation of attitudes towards immigration for each adult child.[12] Table 1.8 reports the number of observations in both samples by exposure to the reform and by the federal state.

Table 1.2 presents descriptive statistics by exposure to the reform. The upper panel reports the summary statistics for *Sample I*, and the lower panel for *Sample II*. Individuals affected by the increase in compulsory schooling are four percentage

[10]The far right-wing parties are Deutsche Volksunion (DVU), Republikaner (REP) and Nationaldemokratische Partei Deutschlands (NPD).

[11]We proxy the state where the mothers attended school by the state of residence of their children.

[12]All adult children in the estimation sample were cohorts subject to nine years of compulsory schooling.

points less likely to report very concerned attitudes about immigration compared to individuals not affected by the reform. Furthermore, they are, on average, 13 to 14 years younger and have around one more year of schooling. The corresponding difference in concerns about immigration between adult children of mothers exposed to and not exposed to the reform is roughly five percentage points (see *Sample II*).

Figures 1-1 and 1-2 show the effect of the formal extension on school attainment. The figures depict average years of schooling against the distance in years to compulsory schooling reform in each state. Since the introduction of the ninth year was staggered across federal states, the first birth cohort affected differs between the states. Both figures illustrate a distinct jump in average years of schooling around the cut-off point.[13] The figures provide initial descriptive evidence that the compulsory years of schooling reform was effective in increasing average years of schooling for all cohorts affected and for treated mothers.

1.5 Research Design

1.5.1 Empirical Strategy

We start by estimating ordinary least squares (OLS) regression with a dichotomous dependent variable for our *Sample I*. The linear probability model linking immigration attitudes to years of schooling takes the form:

$$C = \tau S + \kappa X + State + Cohort + Syear + Trend + \epsilon \tag{1.1}$$

The outcome C equals one if the individual is very concerned about immigration, and zero otherwise (somewhat concerned, or not concerned at all). The variable S is the years of schooling and is assigned based on the school leaving qualification.

[13]Figure 1-2 also reveals an upward sloping trend in average maternal schooling even before the reform, which indicates an increasing educational attainment over the cohorts.

State, *Cohort*, *Syear*, are federal state, cohort, and survey year fixed effects, respectively. Federal state fixed effects control for time-invariant differences in concerns about immigration between the states. They also capture the differential impact of the war (for instance, destroyed housing stock (Braun and Kvasnicka, 2014)) and time-invariant educational differences between states. Cohort fixed effects capture nationwide secular changes in educational attainment over birth years. Since we pool 15 interview years from 1999 to 2013, survey year fixed effects are important for accounting for systematic discrepancies that are borne by political and economic developments in a specific interview year. For instance, events such as 9/11 in 2001, a high unemployment rate in Germany in 2005, or the financial crisis in 2008 may affect individual attitudes. *Trend* is a state-specific linear cohort trend and the vector X includes a male dummy and a quadratic in age as additional controls.[14]

A simple OLS regression is likely to estimate a biased effect of schooling on immigration attitudes. To overcome endogeneity, we instrument years of schooling of an individual with the compulsory schooling reform in West Germany in an IV framework. The first and second stages are specified as follows:

$$S = \delta R + \lambda X + State + Cohort + Syear + Trend + \upsilon \tag{1.2}$$

$$C = \theta \hat{S} + \gamma X + State + Cohort + Syear + Trend + \eta \tag{1.3}$$

where $\hat{S}$ captures the predicted years of schooling, and R is the dichotomous instrumental variable, which equals one if the individual is affected by the reform, and zero otherwise. The remaining variables are defined as in equation (1.1). To obtain accurate statistical inference, we cluster the standard errors over the federal

[14]The state-specific cohort trend is an interaction term between the state dummy and the year of birth. Note that age is the linear combination of year of birth and survey year, hence it is implicitly controlled for.

state-year of birth cell, which results in 296 clusters in our main specification.[15]

1.5.2 Identifying Assumptions

The extension of compulsory schooling from eight to nine years is likely to be binding for students who otherwise would have completed their education after eight years of school attendance, meaning students in the lower school track. In our IV framework, we therefore estimate a local average treatment effect for the lower end of the educational distribution. Compliers in the sample are students who attend school for an additional year due to the compulsory schooling reform. Four assumptions need to be fulfilled to estimate the causal effect of one additional year of schooling on compliers' attitudes towards immigration—namely, the assumption of instrument independence, exclusion restriction, instrument relevance, and monotonicity (Angrist and Pischke, 2009, p. 114).

Instrument independence. The independence assumption requires the instrument (the compulsory schooling reform) to be independent of the vector of potential outcomes and potential assignments to one additional year of schooling. Similar to the previous literature using increases in compulsory years of schooling as a source of exogenous variation, one possible threat to our identification strategy is the potential bias due to policy preferences. For instance, the independence assumption would fail if the compulsory schooling reform was implemented to meet the political agenda of increasing social cohesion, or if the timing of the reform was correlated with immigration attitudes in a given state.

To assure that the independence assumption holds, in addition to federal state and cohort fixed effects, we include state-specific cohort trends. Moreover, we run additional tests. First, Table 1.10 presents a short overview of the curriculum of the

[15] To test the robustness of the inference, we also conducted the analysis with two-way clustering on the dimension of federal state and year of birth. The resulting standard errors are smaller. However, it should be noted that we only have ten clusters at the federal state level, which is susceptible to a "too few" clusters problem (e.g., Cameron and Miller, 2015).

9th year of schooling and reasons for introducing this additional year in the various federal states. For instance, the reasons for and aims of introducing the 9th year in North Rhine-Westphalia were (1) providing a more in-depth general education; (2) reinforcing political education, and (3) practicing, acquiring and expanding basic skills and knowledge. Overall, across all federal states in Germany, the main reason for the compulsory schooling reform was to improve students' occupational maturity (Petzold, 1981). None of the laws and of the relevant literature mention aspects related to immigration, immigration policies, or civic attitudes as relevant aspects. Second, in our robustness section, we show that potential educational drivers such as the number of teachers and their age composition remained largely stable around the reform period in the seven out of ten West German states for which we have reliable data. We also find no evidence of significant changes in immigration flows at the time when the reform was enacted in all the different federal states. Last, in Table 1.11, we present results from an OLS regression with the timing of the reform regressed on pre-reform state characteristics. Neither the student-teacher ratio in the basic track, a group of socio-economic state characteristics such as the share of migrants nor a variable describing the political orientation in a state predict the reform indicator.

Exclusion restriction. The exclusion assumption means that the compulsory schooling reform has no direct effect on concerns about immigration and influences attitudes towards immigration only through the additional year of education. While it is unlikely that perceived concerns about immigration change as a result of mere reform eligibility, the exclusion restriction would fail if compulsory schooling laws systematically altered the inflow or outflow of immigrants to a specific federal state. A higher or lower presence of immigrants due to the reform may affect locals' attitudes towards immigration *per se* or through a change in the political environment. In the robustness section, we address this issue by controlling for the proportion of foreigners during childhood at the federal state level.

Instrument relevance. With respect to the relevance assumption, the compulsory schooling reform must be sufficiently partially correlated with years of schooling, after controlling for the set of exogenous covariates. Otherwise, a weak instruments problem may cause our IV estimation to yield an even larger bias of the estimates than under OLS. The relevance requirement is empirically testable. For a single endogenous variable, the conventional rule suggests that the first-stage F-statistic of a significance test on the excluded instrument takes at least the value of ten (Staiger and Stock, 1997).

Monotonicity. Finally, the monotonicity assumption implies that while the instrument might have no effect for a certain proportion of the population, all those who are affected are affected in the same direction (Angrist and Pischke, 2009, p. 114). In the context of compulsory schooling, this implies that no student reduces his or her educational attainment because of the introduction of one additional year of schooling, i.e. there are no defiers. Although the monotonicity assumption is not directly verifiable, Angrist and Imbens (1995) suggest that it has a testable implication. They suggest plotting the cumulative distribution functions (CDF) of the treatment variable by the binary instrument. The monotonity assumption is likely to hold if the two CDFs do not cross.

In Figure 1-3, we plot the CDFs of individual schooling years by exposure to the compulsory schooling reform (affected or not affected). The figure shows that the CDFs for individuals affected by the reform lies below the CDF for individuals not affected by the reform (stochastic dominance). The two CDFs do not cross. To further assure that the reform did not affect the track choice of individuals, we estimate reduced form regressions with the track choice as an outcome, and find that the introduction of the ninth class does not have a significant effect on the choice of the school track (see Table 1.9 in the appendix).

1.6 Empirical Evidence

1.6.1 Own Schooling and Immigration Attitudes

Table 1.3 shows the effect one additional year of schooling has on concerns about immigration. The focus is on *Sample I*, i.e., the relationship between an individual's own education and attitudes towards immigration. Columns (1) and (2) display OLS estimates, while columns (3) and (4) describe the results for the IV framework. Moreover, models in columns (1) to (4) include controls for sex, a quadratic in age, as well as a full set of federal state, birth cohort, and survey year fixed effects. To mitigate concerns that the introduction of the reform within states may be correlated with trends in education and immigration attitudes, in columns (2) and (4), we also control for linear state-specific cohort trends.

What unifies all findings in Table 1.3 is the clear-cut negative correlation between schooling and high concerns about immigration. OLS estimates in columns (1) and (2) suggest that one additional year of schooling is associated with a decrease of 4.4 percentage points (14.6 percent) in the probability of having high concerns about immigration. However, OLS estimates in the present setting may not convey the unbiased effect of schooling. Hence, we turn to the IV estimates.

Panel B of Table 1.3 reports the first-stage results. The first-stage estimates in columns (3) and (4) suggest that the compulsory schooling reform increases years of schooling by around 0.54 years. The point estimates are consistent with the findings of Pischke and von Wachter (2005) and Siedler (2010), who study the same reform in different contexts with different data-sets.[16] In both specifications, the first-stage F-statistics of the instrument are much larger than ten. Hence, the null hypothesis that the coefficient on the reform—δ in equation (1.2)—is equal to zero can easily be rejected (Staiger and Stock, 1997).

[16]To be precise, Pischke and von Wachter (2005) find that the reform increased years of schooling by 0.59 years when using data from the Qualification and Career Survey and 0.55 when using data from the Micro Census. Siedler (2010) finds that the effect of the reform ranges from 0.39 to 0.54 when using ALLBUS data and from 0.43 to 0.52 when using ForsaBus data.

IV estimates in Panel A of Table 1.3 confirm that schooling has a negative effect on the probability of reporting high concerns about immigration. The inclusion of state-specific cohort trends only slightly alters the magnitude of the effect. In our preferred specification in column (4), an additional year of schooling decreases the probability of being very concerned about immigration by 6.3 percentage points. The effect is statistically significant at the five percent level. Against the base level of 30 percent of the full estimation sample, this translates into a decrease of around 20 percent, which is a considerable effect.[17] When estimating our main regression separately by gender, the effect of schooling is larger for women (-0.08 (0.047)), however the estimates are not statistically different across sexes.

Are the estimated IV returns to schooling plausible? In an attempt to answer this question, we compare our findings with a related study. Closest to our study is d'Hombres and Nunziata (2016) who estimate education effects on pro-immigration attitudes averaged over a pool of 12 European countries. The magnitude of our IV estimates is comparable to the findings of d'Hombres and Nunziata (2016). They find point estimates that are in the range of 6 to 11 percentage points.[18]

We also consult publications and historical documents to shed light on the curriculum of the ninth compulsory year of schooling. First, it should be noted that the same content was not merely spread over more schooling years (Leschinsky and Roeder, 1980). Rather, the main focus of the additional school year was on civic and occupational education. An official document published by the state parliament of North-Rhine Westphalia reads: "There is also complete agreement that political education, the foundations of which can only be partially laid during the first eight years of school, must be a key component of overall educational work in the ninth school year." (Nordrhein-Westfalen, 1962, p. 62, own translation).

[17]In unreported regressions, we also estimated the reduced form—the direct effect of compulsory schooling reforms on attitudes towards immigration. The direct effect is three percentage points and is statistically significant at the ten percent level.

[18]Although the authors do not report the magnitudes of the effect in terms of percent, we attempted to recover these on the basis of the summary statistics reported on page 205 in their study. The effects range from 15 to 23 percent.

Similarly, with respect to the additional ninth year, Leschinsky and Roeder (1980) write: "The focus, however, was on expanding certain areas of learning, where the focus was on introducing students to their role as citizens, and on career and labor market orientation." (Leschinsky and Roeder, 1980, p. 334, own translation). Table 1.10 in the appendix summarizes the key goals of the additional ninth school year for the different federal states. It documents that the ninth year is mainly aimed at general education, rather than teaching specific skills.

Overall, the discussion indicates that the negative instrumental variable estimate of an additional year of schooling on high concerns about immigration shown in Table 1.3 appears to be quite plausible.

1.6.2 Maternal Schooling and Offspring's Immigration Attitudes

Immigration attitudes may be shaped not only by an individual's own educational attainment but also by parental education. The relationship between parents' education and the immigration attitudes of their offspring is likely to be multi-fold. On the one hand, the economic literature has repeatedly confirmed the intergenerational transmission of education (Plug, 2004; Black et al., 2005; Holmlund et al., 2011; Piopiunik, 2014; Dickson et al., 2016). On the other hand, young adults, whose parents were very concerned about immigration to Germany during their childhood years, are more likely to express strong concerns about immigration themselves (Avdeenko and Siedler, 2017).

Little is known about the effect of parental education on non-educational outcomes of their children.[19] In particular, to the best of our knowledge, intergenerational effects of schooling on adult children's attitudes and values have not been studied to date. We now examine whether a mother's education affects her

[19] As an example, exceptions are Currie and Moretti (2003) for birth outcomes of children, Kemptner and Marcus (2013) for health outcomes, and Lundborg et al. (2014) for cognitive and non-cognitive skills.

offspring's attitudes towards immigration. The relevant first- and second-stage equations are:

$$S^m = \beta R^m + \delta X^o + State^m + Cohort^m + Syear^o + Trend^m + \zeta \tag{1.4}$$

$$C^o = \alpha \hat{S}^m + \omega X^o + State^m + Cohort^m + Syear^o + Trend^m + \nu \tag{1.5}$$

In these equations, the superscript o denotes the offspring and m the mother. C^o is the dichotomous measure of immigration attitude of the adult child and is defined as in the main analysis. The variable $\hat{S^m}$ refers to predicted years of maternal schooling, R^m is a dichotomous variable indicating whether the mother was exposed to the reform. $State^m$, $Cohort^m$, and $Syear^o$ are state, cohort and survey year fixed effects, respectively, and $Trend^m$ is a vector of state-specific cohort trends. X^o controls for the sex of the child. The sample includes all mothers born between 1930 to 1960 (*Sample II*).

Table 1.4 reports the results. In all specifications maternal education has the expected sign. The higher the mother's schooling, the less likely the adult children are to report very concerned attitudes towards immigration. The OLS estimates point to a negative association of roughly three percentage points. The second stage IV estimates suggest that an additional year of maternal schooling reduces the probability of the mother's offspring reporting very concerned attitudes about immigration by 8.3 to 6.8 percentage points. As previously, we prefer the specification with state-specific cohort trends (column 4).[20]

The effect of a mother's schooling on her offspring's attitude is comparable to the main effect of an individual's own schooling on his or her own attitudes towards immigration (-6.8 and -6.3 percentage points).[21] A likely explanation for

[20]The coefficient on mothers education remains precisely estimated and of comparable magnitude also after controlling for offspring's own education.

[21]If we also control for father's schooling, the estimate on mother's schooling is virtually the same (with standard errors slightly larger), while the coefficient on father's schooling itself is

the comparable magnitudes is that exposure to the compulsory schooling reform has both educational and attitudinal spillovers: while the offspring of these mothers are likely to obtain a higher level of education themselves (Piopiunik, 2014), they are also likely to partially inherit the attitudes of their mothers (Avdeenko and Siedler, 2017).[22] It should be noted that in the first-stage estimations (Panel B, Table 1.4), we find a larger effect of the compulsory schooling reform on maternal schooling compared to the first-stage estimates of the instrument on own education in Table 1.3. This finding is anticipated, as the average number of years of schooling among women in these cohorts is considerably lower than that of men. As a result, the average increase in years of schooling due to the reform is greater.

Overall, our findings suggest a strong effect of schooling on immigration attitudes, directly and through maternal education. The next subsection tests the robustness of the findings. Whenever possible, we report the robustness checks for both *Sample I* and *Sample II*.

1.6.3 Robustness Checks

Teachers' age and gender composition. An alternative explanation for the effects may be environmental changes in schools that coincided with extended schooling. For instance, consider that the extension coincided with (or led to) hiring more teachers, in particular young ones, who may have more liberal attitudes, and these teachers may pass on their positive attitudes to their students. To address this possibility, we collected data on teachers' age in basic track schools. Figure 1-4 in the appendix plots the number of teachers in the basic school track in

positive, but insignificant. Including father's schooling enables us to obtain the effect of the mother's schooling on her offspring's attitudes, net of assortative mating. Note, however, that schooling of spouses is highly collinear, thus inflating the standard errors and complicating the *ceteris paribus* interpretation of the coefficient.

[22] In unreported regressions, we use the self-reported information by the mothers themselves and merge them with their children. The resulting sample is much smaller in size. The OLS point estimates are highly comparable to those in Table 1.4. IV point estimates continue to be negative but are imprecisely estimated.

different age categories. The vertical lines denote the reform date in each state. We are only able to construct these figures for seven federal states, because the reports by the Federal Statistical Office are inconsistent prior to 1958. The plots do not reveal a discontinuous change in the number of young teachers around the cut-off point. Furthermore, reflecting upon the accumulating evidence on the gender gap in political preferences (Edlund and Pande, 2002; Box-Steffensmeier et al., 2004), Figure 1-5 plots the share of female teachers in basic track over 1950-1970. The figures show that the share of female teachers has been rising from the mid 1950s onward, however, there are no discontinuous changes around the reform dates. This suggests that the impact of schooling on high concerns about immigration is unlikely to be driven by a sudden increase in young and more tolerant teachers.

Geographic mobility. In *Sample I*, the information on the federal state where the individual attended school, is available for half of our estimation sample. Similar to Pischke and von Wachter (2008) and Siedler (2010), we use the federal state of residence to proxy the school attendance state if the state of school attendance is unknown. If geographic mobility is high between federal states with different dates of reform introduction, this raises concerns about potentially imprecise assignment of the instrument. Descriptive evidence suggests that for 81 percent of individuals, the federal state of residence and the federal state of school attendance coincide. When we also consider that four federal states introduced the reform simultaneously, potential misspecification of the relevant reform date drops to 13.6 percent. Note that these figures are likely to be an upper bound, as the question on federal state of school attendance is included for younger cohorts in the SOEP, and older cohorts were geographically less mobile.

To further examine the issue of geographic mobility, we restrict the analysis to different sub-samples. In Table 1.5, we present results for individuals with valid information on the state of school attendance (specification 1) and current state of residence (specification 2), respectively.[23] Thus, we only estimate the regressions

[23] We are only able to carry this analysis for *Sample I* due to data limitations.

for individuals for whom we have information on their state of school attendance and their current state of residence. Assignment of the instrument based on state of school attendance generates a stronger first-stage, as a pairwise comparison of the first-stage F-statistics in specification 1 and 2 of Table 1.5 reveals (130.8 compared to 79.9). However, IV estimates are highly comparable to each other. In line with Pischke and von Wachter (2008) and Siedler (2010), this suggests that proxying the state of school attendance with the current state of residence is unlikely to be problematic.

Leaving out the first affected cohort. The four states, which introduced the ninth class in 1966-67, also introduced two short school years due to a change in the start of the school year.[24] One short school year began on 1 April, 1966 and ended on 30 November, 1966, and a second short school year started on 1 December, 1966 and ended on 31 July, 1967 (Pischke, 2007). The nominal curriculum did not change for the transition cohorts, and Pischke (2007) further argues that these short school years did not lead to a reduction in human capital accumulation.[25] The determination of the first birth cohorts affected by the increase in compulsory years of schooling reflects the presence of short school years. Nevertheless, in some states, the actual implementation of the compulsory schooling reform might not have perfectly coincided with the official date. For these reasons, we repeat our analysis by excluding the first affected birth cohorts.[26] Row 1 of Table 3.17 shows the results. The estimate on the effect of own schooling is largely unaffected by this modification, while the corresponding one for maternal schooling is 4 percentage

[24] The four federal states are Hesse, North-Rhine Westphalia, Rhineland-Palatinate and Baden-Wuerttemberg.

[25] According to the Agreement on the Unification of the School System (*Hamburger Abkommen*), the start of the school year should have been moved to the end of the summer in all federal states by 1967 (Pischke, 2007). The implementation varied across federal states. West Berlin and Hamburg opted for one long school year. Bavaria already started the academic term in summer. Lower Saxony added an additional school period along with shorter school years in subsequent years for some types of schools. For further details see Pischke (2007).

[26] For instance, Siedler (2010) mentions that some schools in Bremen may have increased compulsory schooling in 1959 rather than in 1958 (the official introduction date).

points larger. This difference may indicate potential measurement error in the assignment of reform exposure for cohorts within one year of the cut-off, if adult children retrieve the birth year of their mothers with a small measurement error.

Region-specific year of birth effects. Stephens and Yang (2014) examine the common trend assumption in studies which use compulsory schooling laws as instruments in their analysis. The authors argue that the assumption of common trends in variables affecting different cohorts across states is unlikely to hold. In our baseline specification, we control for state-specific cohort trends, that should capture within-state developments correlated with the changes in law, educational improvements, and the outcome. Similar developments, however, may also occur at a regional level. Hence, to allow birth cohort effects to vary across regions, we divide West Germany into three broad regions—northern, central, and southern. The regional split is as follows: the northern part includes the states of Hamburg, Bremen, Schleswig-Holstein and Lower Saxony, the central part includes North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Saarland, and the southern part consists of Baden-Wuerttemberg and Bavaria.[27] Including these regions-by-year-of-birth fixed effects results in second-stage IV coefficients on schooling that are roughly three percentage points larger than the coefficients from the baseline specification for both samples (specification 2).[28]

Social desirability bias. Questions on immigration attitudes may be sensitive to social desirability and, consequently, might trigger reporting bias. If the well educated are more likely to identify the "socially desirable" answer in the presence of the interviewer and report it, we risk overestimating the true effect of

[27]Stephens and Yang (2014) use US census regions—North-East, Midwest, South and West. The findings show that for the majority of studies examined, IV estimates become either close to zero or have the “wrong” sign.

[28]In unreported regressions we also added state-specific quadratic trends. The coefficient on education in Sample I remains identical. For Sample II the coefficient on mothers education is somewhat smaller in magnitude and imprecisely estimated. However, the quadratic trends are neither jointly nor individually statistically significant, suggesting that the specification with quadratic trends is the wrong specification.

schooling, unless these respondents also act in a "socially desirable" manner. We are able to partially address this problem since the SOEP follows a "mixed mode approach" in interview methods. In specification 3 of Table 3.17 we control for the interview mode by constructing a dichotomous variable that reflects whether the questionnaire is self-administered. This has virtually no effect on the main estimates, suggesting that social desirability bias is unlikely to be a concern.[29]

Bandwidths around the reform. We restrict the sample to cohorts (mothers) born between 1930 and 1960. While this selection is data efficient, it has its limitations: we create an asymmetric window around the first affected birth cohort in each state. This means that early implementer states contribute more to the treatment group, while late implementers contribute more to the control group. To assure that this sample restriction does not affect the main results, and to further account for potential omitted time-varying changes in attitudes and education, we re-estimate equations (1.2) to (1.3) and (1.4) to (1.5) in a regression discontinuity framework. To this end, for each state, we construct a pre-treatment and post-treatment sample composed of cohorts born six years before and after the reform. We repeat the exercise by providing a slightly broader bandwidth—eight years before and after the reform. This mode of sample restriction creates a symmetric window around the cut-off point in each federal state, with similar sample sizes for treatment and control groups by states. Moreover, it may also reduce concerns about unobserved developments that state-specific linear trends may not capture. However, these restrictions also result in smaller sample sizes compared to the main samples.

Table 3.17 presents the results in rows 4 and 5. The point estimates on own schooling are larger compared to the baseline effect in the order of one to two percentage points. The opposite, however, applies to maternal schooling. The coefficients are smaller in magnitude and imprecisely estimated. The restriction by bandwidth may increase measurement error problems in mothers' year of birth:

[29]For the descriptive statistics of the interview methods used, see Table 1.12 in the appendix.

if retrospective retrieval of maternal year of birth occurs with an error of one year, it would result in distorted assignment of reform exposure immediately before and after the cut-off.

Childhood political and economic environment controls. A common concern in IV estimations is the possibility of omitted variable bias. Changes in the political and economic environment that coincide with the timing of compulsory schooling reform may be a threat to the exclusion restriction. To alleviate these concerns, we re-estimate our main specification by controlling for a range of proxies for political climate of the state when the individual was 15 years old (specification 6). In particular, we add the voter turnout in latest federal elections, share of votes that the two major parties Christian Democratic Union (CDU) and Social Democratic Party (SPD) receive, also for the number of seats each of those parties acquire in the state parliament. In specifications 7 and 8 we control for GDP per capita, share of unemployed individuals in the population, and share of foreigners at the federal state level.[30] Specification 7 shows the results when we include GDP per capita, unemployment rate and migration inflow over ages 12 to 18, while specification 8 controls for the above mentioned variables for the interview year. The point estimates on schooling for both samples are precisely estimated and highly comparable to our IV estimates in Tables 1.3 and 1.4.[31]

Ordered IV probit. In ordered IV probit regressions, we maintain the original three-category coding of the outcome. The IV probit coefficients on schooling are reported in specification 9 of Table 3.17. The estimates are negative and statis-

[30]The data on GDP is provided by the official working group "Arbeitskreis Volkswirtschaftliche Gesamtrechnungen der Länder", the number of foreigners and the population data comes from the Federal Statistical Office of Germany. GDP per capita is measured in Deutschmark for specification 7, and in euros for specification 8. Data on the unemployment rate at the state level is only available as of 1960, so we use unemployment figures from the Federal Employment Agency and adjust these according to the population size.

[31]In unreported regressions, we also control for the residence community type of the respondent. Although choice of residence (urban or rural) may be endogenous, we reassuringly find that our point estimates on schooling are virtually unaltered.

tically significant at the five percent level, confirming that schooling negatively affects immigration concerns. The average marginal effects suggest that an additional year of schooling reduces the probability of reporting "very concerned" attitudes by 4.9 percentage points, with a standard error of 0.024 for individual's own schooling, and by 5.1 percentage points with a standard error of 0.021 for the maternal schooling.[32]

Average concerns over all available panel years. Despite the appealing panel structure of the SOEP, fixed effect estimation is not feasible in our analysis. The reason for this is that the main explanatory variable of interest—years of schooling—is time-invariant within individuals. Pooling all available panel years and treating it as a pseudo cross-section gives more weight to individuals who stay in the panel survey longer, and this pattern may be correlated with the schooling reform. Additionally, Lancee and Sarrasin (2015) show that there is little variation in immigration attitudes within individuals. Therefore, in our main analysis, we use the first available observation for each individual. To reassure that this sample selection criteria does not drive the results, we report the estimates with an alternative definition of the dependent variable in specification 10 of Table 3.17. We create a single measure for each individual by calculating the proportion of times that a person reported being very concerned about immigration over all available panel years. The dependent variable is now continuous and ranges from zero to one.[33] The resulting coefficient on schooling is comparable with the findings in the main section. An additional year of schooling reduces the probability of being very concerned about immigration by 4.3 and 5.4 percentage points in *Sample I* and *Sample II*, respectively, which is a reduction of 15 percent

[32]The marginal effects on reporting "somewhat concerned" or "not concerned at all" are 0.006 (0.003) and 0.044 (0.021), respectively, for *Sample I* and -0.006 (0.002) and 0.057 (0.222) for *Sample II*.

[33]Note that this specification includes fixed effects for the number of times the respondent replied to the question, and average age along panel years for the individual's own schooling sample.

and 23 percent against the base level of average concerns.

Placebo tests. In specification 11 of Table 3.17 we randomly generate reform implementation dates from a uniform distribution for each state over the time-span of 1949-1969, and hence "fake" reform exposure. The first-stage F-statistics for both *Sample I* and *Sample II* are very small. Accordingly, the estimates in the second stage are imprecise and implausible in magnitudes.[34] As an additional placebo test (not reported), we run our main specification with placebo outcomes, namely parental schooling of the individuals (*Sample I*). Reassuringly, the coefficient on schooling is statistically insignificant.

1.7 Channels

1.7.1 Channels of Own Schooling

From a policy perspective, distinguishing the channels is necessary to properly address concerns about immigration. If attitudes are driven by a more secure labor market position, then policy measures should target labor market institutions, while if attitudes stem from deeper knowledge or rectified reasoning, then policy measures should target quality of education. We start by considering potential labor market outcomes through which education might affect individual attitudes towards immigration and extend our analysis to generalized trust as a proxy for social capital.

Labor market outcomes. According to the canonical model, if a country experiences low-skilled migration influx (hence an increase in low-skilled labor supply), the wages for the native population in a comparable skill group will decrease, whereas high-skilled natives might benefit (Dustmann et al., 2016). Hence, if attitudes towards immigration are motivated through fear of labor market competi-

[34] In unreported regressions we also use the fake reform in a reduced form, and find that the fake reform has no significant effect on immigration attitudes in either of the samples.

tion, in countries with low-skilled immigration the highly skilled natives should be less likely to have negative attitudes towards immigration. However, the empirical literature is not univocal on the direction and magnitude of the impact immigration has on labor market outcomes of the natives.[35] The findings from the literature on immigration attitudes do not provide unambiguous empirical support for the labor-market-competition hypothesis either (e.g. Scheve and Slaughter, 2001; Mayda, 2006; Hainmueller and Hiscox, 2007; Card et al., 2012; Hainmueller et al., 2015).

To test whether labor market outcomes are behind our findings we examine a range of variables that should capture job market competition fears at different margins. We use our main identification strategy to test whether the compulsory schooling has a significant effect on the hypothesized channel. We start at the extensive margin by analyzing whether compulsory schooling affects labor force participation and employment status at the time of the survey. To encompass larger time-span we also examine the probability of ever being unemployed until the time of the survey (row 1-3, Table 1.7). The point estimates suggest that an additional year of education increases the chances of participating in the labor force, and decreases the probability of being either contemporaneously or ever unemployed. However, none of the estimates are statistically different from zero.

Our focus is next on occupational outcomes, in particular the likelihood of being in a white-collar occupation, the likelihood of being employed as a civil servant and the likelihood of having an unlimited work contract.[36] The interest in the white- and blue-collar binary is motivated by the idea that the ones in blue-

[35]Dustmann et al. (2016) show that the parameters from different models are often not comparable and differ in their interpretation. In particular the national skill-cell approach identifies an effect for different experience levels within education groups, while the pure spatial approach estimates the total wage effect on a particular skill group. Furthermore, the authors argue that the national skill-cell approach implicitly assumes that natives and immigrants in the same education-experience cell are perfect substitutes, which introduces bias in the parameters.

[36]White-collar occupation is defined as occupational codes 1-5 based on the ISCO-88 occupational classification. This definition approximately translates into non-manual and manual classification as well.

collar occupations are more likely to directly compete with immigrants in the job market, given that historically the immigration to Germany has been low-skilled. On the other hand having an unlimited contract or being a civil servant provides the highest level of job security. Hence, these subgroups should be least worried as far as job market competition is concerned. However, again we find no evidence that education significantly changes the likelihood of these outcomes.

Moreover, we consider the effect of additional education on individual's net monthly labor earnings and monthly household income hypothesizing that higher financial security may decrease the probability of perceiving immigration as an economic threat. Again, the point estimates indicate the expected direction, however they are statistically insignificant. Using the same compulsory schooling reform, Pischke and von Wachter (2008) find zero returns to compulsory schooling in terms of wages and speculate that this is because in Germany students generally learn skills relevant for the labor market earlier than in other countries. These findings are confirmed by Kamhöfer and Schmitz (2016) using SOEP data. Furthermore, labor market outcomes do not appear to be significantly correlated with immigration attitudes, after controlling for schooling (see Panel A of Table 1.13).

Worries about immigration can be motivated through real as well as perceived threat. And even though the findings in this section suggest no statistically detectable schooling effect on realized labor market outcomes, it might still have an impact on how individuals perceive their own economic situation. To capture this dimension we additionally examine concerns about job security and own economic situation in Panel B of Table 1.7.[37] Again, although education reduces the probability of being very concerned along these two dimensions, the effects are not statistically different from zero.

Social capital. Putnam defines social capital as "features of social organization

[37]Concerns about job security and own economic situation have three response-categories: "very concerned", "somewhat concerned", and "not concerned at all". We recode the original three-category variables into a dichotomous variable that equals one for individuals who report that they are "very concerned", and zero otherwise.

such as networks, norms, and social trust that facilitate coordination and cooperation" (Putnam, 1995, p. 67). While social capital has long been linked to economic growth and to the quality of institutions (Knack and Keefer, 1997; Zak and Knack, 2001), a recent literature proposes that it is also strongly linked to attitudes towards immigration. Herreros and Criado (2009) argue that immigration concerns are often a consequence of fears of the out-group. Social trusters avoid racial and cultural stereotyping when forming their beliefs about others' trustworthiness, hence they extend their trust also to foreigners. In a cross-country analysis of 19 European countries, the authors find that higher levels of social trust are associated with positive attitudes towards immigration. Economidou et al. (2017) and Halapuu et al. (2013) also confirm that across Europe individuals with high social capital exhibit a more positive attitude toward immigration.

Here we focus on one, yet crucial dimension of social capital—social trust. Delhey and Newton (2005) define it as "the belief that others will not deliberately or knowingly do us harm, if they can avoid it, and will look after our interests, if this is possible" [p. 311]. The determinants of trust are not well understood. Nevertheless, education is posited to be one of the primary determinants (Helliwell and Putnam, 1999; Alesina and La Ferrara, 2002). In a meta analysis of 28 studies Huang et al. (2009) conclude that an additional year of schooling increases social trust by 4.6 percent of its standard deviation.[38]

We examine whether compulsory schooling affects general trust. The SOEP has measures on trust in 2005, 2008 and 2013. Respondents are asked to evaluate the following three statements on a four-point scale: (1) on the whole one can trust other people, (2) nowadays one cannot rely on anyone, (3) if one is dealing with strangers, it is better to be careful before one can trust them.[39] The validity of SOEP trust measures is reported in Fehr et al. (2003).

We first generate the mean of each item for each person over all available

[38]Note, however, that the majority of the studies examined do not take into account the endogeneity issue of education.

[39]The scale has the following items: strongly agree, agree, disagree, and strongly disagree.

years. We subsequently collapse the respective items into a single standardized measure of trust with a mean of zero and standard deviation of one, and then link our estimation sample to trust outcomes.[40] Panel B of Table 1.7 shows the impact of an additional year of schooling on social (general) trust. An additional year of schooling increases trust by 17.6 percent of its standard deviation. This effect is precisely estimated at the five percent significance level. Furthermore, as Table 1.13 shows, higher trust is strongly correlated with less concerns about immigration, after controlling for schooling.

Our analysis in this section suggests that labor market outcomes are unlikely to be behind the finding that education reduces the likelihood of high concerns about immigration. It indicates, however, that returns to education in terms of social trust are potentially mediating the effect of education on attitudes towards immigration.

1.7.2 Channels of Maternal Schooling

The channels through which a mother's schooling affects her children's attitudes are likely to be multifaceted. Below we discuss the potential mechanisms in the light of the previous literature. Due to data limitations, however, we are unable to disentangle the role and importance of each mechanism.

In theories on the origins of preference formation, parents transmit their attitudes and preferences to their children, actively or passively, leading to persistences across generations (Bisin and Verdier, 2001). Related to this, Avdeenko and Siedler (2017) find that young adults, whose parents were very concerned about immigration to Germany during their childhood years, have a much higher likelihood of also expressing strong concerns about immigration. The marginal effects for the intergenerational transmission in attitudes towards immigration vary between 23

[40]Note that alternatively we can either restrict our main estimation sample to the outcomes from the years 2005, 2008 and 2013 for trust, or take one of the measures of the individual as constant and link it to the main estimation sample. The results are virtually unaffected by either approach.

to 30 percentage points, depending on the specification and parent-child gender pairs. As we show, cohorts exposed to the compulsory schooling reform are less likely to be worried about immigration. Consequently, building on the findings of Avdeenko and Siedler (2017), their offspring should also be less likely to be worried about immigration.

Another relevant channel through which a mother's schooling may have an impact on her children's immigration attitudes is through an induced increase in her children's educational attainment. For our estimation sample, we find a positive, yet statistically insignificant effect of maternal schooling on the educational attainment of the mother's offspring. However, exploiting the same compulsory schooling reform, with somewhat different sample restrictions, Piopiunik (2014) finds that an additional year of schooling for the mother significantly increases the likelihood of her sons completing intermediate or higher track schooling (10 years or more). The effect does not hold for daughters.

We also analyze whether mother's schooling affects her offspring's social trust. Unreported findings suggest a positive and insignificant effect on social trust. Although we do not find any evidence that maternal schooling increases the trust of adult children, the previous section documents a positive significant effect of an individual's own schooling on social trust. Consequently, the higher level of maternal trust due to the mother's additional year of schooling may have spillover effects on the trust level of her offspring. Related to that, Dohmen et al. (2012) analyze the intergenerational correlation in trust using the SOEP and find that increasing the mother's trust by one standard deviation increases the child's trust by around 0.24.

Overall, the previous literature provides suggestive evidence that a reduction of adult children's strong concerns about immigration due to higher maternal education might be driven by one or a combination of multiple mechanisms: a positive effect of maternal education on a child's educational attainment, which in turn changes adult child's attitudes towards immigration; an effect of the mother's education on the child's attitudes through modifying her own attitudes; a positive effect of maternal trust on child's trust, which in turn changes the adult child's attitudes towards immigration. We should reiterate that, due to data limitations, we are not able to disentangle the role and importance of each mechanism.

1.8 Conclusion

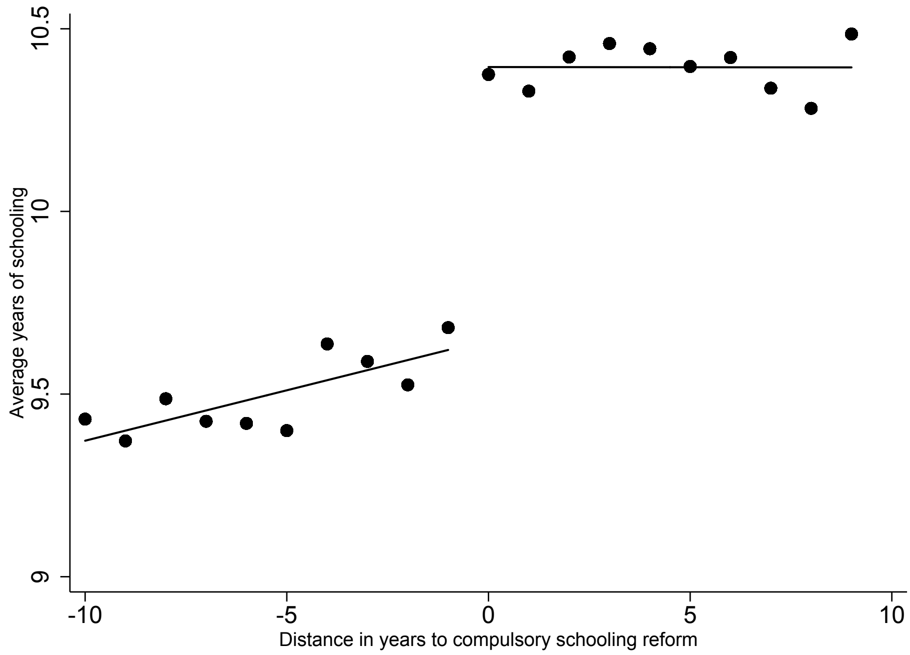
The present study examines the role of education in shaping immigration attitudes. Although previous literature establishes a strong positive empirical correlation between education and pro-immigration attitudes, this association cannot be interpreted as causal. Simple OLS regressions may produce inconsistent estimates of education due to selection on unobservables. We use compulsory schooling reforms in West Germany as a source of exogenous variation to estimate the consistent effect of schooling on immigration attitudes. Our results show that schooling has a sizable impact on attitudes towards immigration. An additional year of schooling reduces the likelihood of being very concerned about immigration by around 20 percent. We extend the analysis to potential intergenerational transmission of a mother's education to her offspring's immigration attitudes. Our results imply that returns to compulsory schooling extend to the next generation. Adult children's high concerns about immigration decrease by almost seven percentage points with an extra year of maternal compulsory schooling, suggesting a composite effect of both educational and attitudinal spillovers. We show that these findings are robust to alternative specifications and sample selection criteria.

To shed some light on the mechanisms behind our estimates, we analyze potential channels through which an individual's own schooling may affect attitudes towards immigration. The findings suggest that labor market outcomes are unlikely to be behind our results, since there is no evidence on the effect of schooling on considered labor market outcomes. However, the evidence indicates that schooling significantly increases an individual's social capital.

Our empirical findings indicate that education may be an important tool to address concerns about immigration in a host country. Intergenerational transmission of educational attainment and attitudes from parents to their offspring implies that educational policies may address social cohesion through (at least) two generations. The role of education appears to expand over providing financial and labor market welfare, by seeding and molding tolerant beliefs and attitudes.

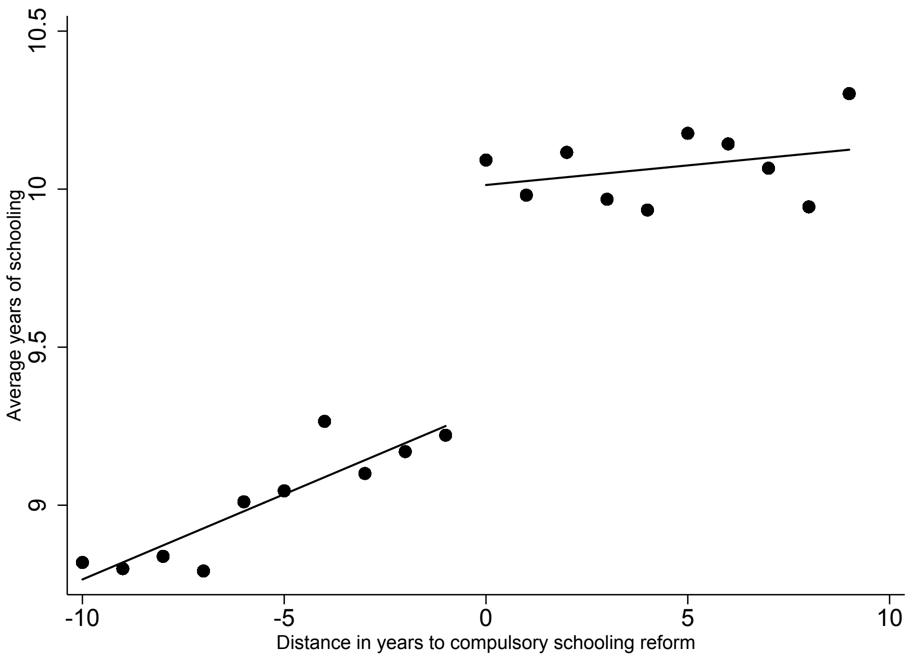
1.9 Figures and Tables

Figure 1-1: The effect of the reform on average years of schooling; *Sample I*



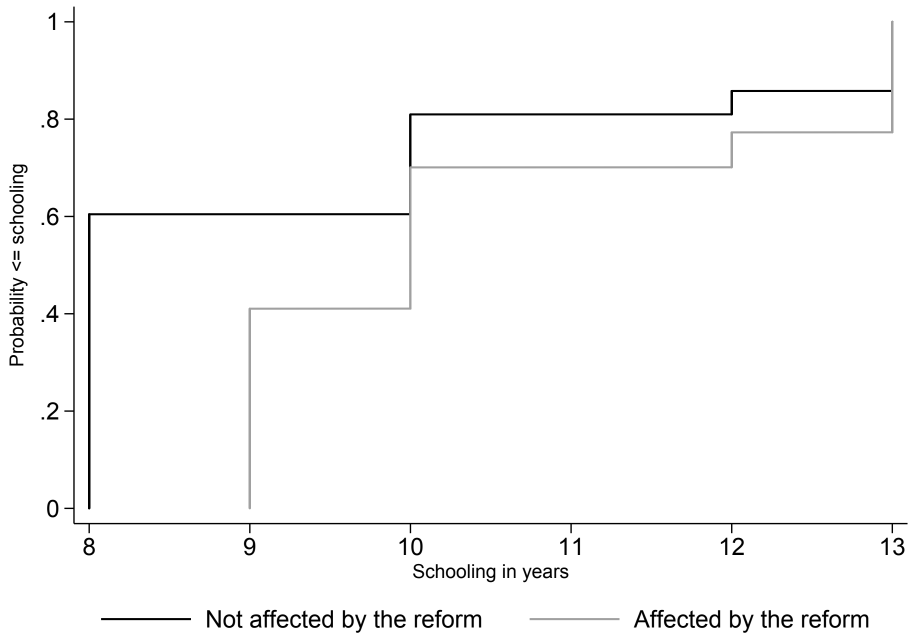
Source: Socio-Economic Panel (SOEP), own calculations.

Figure 1-2: The effect of the reform on average years of maternal schooling; *Sample II*



Source: Socio-Economic Panel (SOEP), own calculations.

Figure 1-3: Cumulative distribution function of schooling by reform exposure



Source: Socio-Economic Panel (SOEP), own calculations.

Table 1.1: Introduction of the ninth class in the basic track of secondary schooling

State	Year of reform	First affected birth cohort
Hamburg	1949	1934
Schleswig-Holstein	1956	1941
Bremen	1958	1943
Lower Saxony	1962	1947
Saarland	1964	1949
Baden-Wuerttemberg	1967	1953
Hesse	1967	1953
North Rhine-Westphalia	1967	1953
Rhineland-Palatinate	1967	1953
Bavaria	1969	1955

Source: Pischke and von Wachter (2005). Federal states are sorted in ascending order by year of reform.

Table 1.2: Descriptive statistics

	Not affected by reform		**Affected by reform**	
	Mean	Std. dev.	Mean	Std. dev.
Sample I:				
Very concerned about immigration	0.317	0.465	0.275	0.447
Years of schooling	9.314	1.849	10.415	1.604
Male	0.495	0.5	0.487	0.5
Year of birth	1941.8	6.456	1955.1	4.577
Age	61.59	7.975	48.07	6.781
Number of observations	8,388		4,810	
Sample II:				
Child very concerned about immigration	0.264	0.44	0.217	0.412
Mother's years of schooling	8.756	1.405	10.097	1.422
Male child	0.483	0.5	0.485	0.5
Mother's year of birth	1940.4	6.651	1954.8	4.704
Child's year of birth	1967.0	8.090	1981.5	6.819
Child's age	36.07	9.356	23.05	6.824
Number of observations	8,628		3,542	

Notes: Concerns about immigration are coded as dichotomous variables, taking on the value one if the individual is very concerned and zero if somewhat concerned or not concerned at all.

Table 1.3: Schooling and immigration attitudes

Sample I	OLS		IV	
	(1)	(2)	(3)	(4)
	Panel A: OLS and second-stage			
Schooling	-0.044^{***}	-0.044^{***}	-0.060^{*}	-0.063^{**}
	(0.002)	(0.002)	(0.031)	(0.031)
	Panel B: First-stage			
Reform			0.545^{***}	0.539^{***}
			(0.050)	(0.051)
F-statistic			116.8	110.3
Observations	13,198	13,198	13,198	13,198
Federal state fixed effects	Yes	Yes	Yes	Yes
Survey year fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes
State-specific cohort trend	No	Yes	No	Yes

Notes: In Panel A, the outcome is a dichotomous variable which takes on the value one if the individual is very concerned about immigration and zero otherwise. In Panel B, the outcome is years of schooling. All regressions also include a male dummy and quadratic in age as controls. Each column represents the coefficient from a different regression. Robust standard errors are clustered at the state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.4: Maternal schooling and offspring's immigration attitudes

Sample II	OLS		IV	
	(1)	(2)	(3)	(4)
	Panel A: OLS and second-stage			
Maternal schooling	-0.029^{***}	-0.028^{***}	-0.083^{**}	-0.068^{**}
	(0.002)	(0.002)	(0.034)	(0.032)
	Panel B: First-stage			
Reform			0.681^{***}	0.684^{***}
			(0.075)	(0.078)
F-statistic			82.4	77.0
Observations	12,170	12,170	12,170	12,170
Federal state fixed effects	Yes	Yes	Yes	Yes
Survey year fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes
State-specific cohort trend	No	Yes	No	Yes

Notes: In Panel A, the outcome is a dichotomous variable which takes on the value one if the adult child is very concerned about immigration and zero otherwise. In Panel B, the outcome is maternal years of schooling. All regressions also include a male dummy for the offspring. Each column represents the coefficient from a different regression. Robust standard errors are clustered at the state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.5: Robustness check: Geographic mobility

Sample I	Schooling	F-stat
1. Last state of school attendance	-0.083^{**}	130.8
	(0.032)	
	[7,440]	
2. Current state of residence	-0.089^{**}	79.9
	(0.039)	
	[7,440]	

Notes: The outcome is a dichotomous variable which takes on the value one if the individual is very concerned about immigration and zero otherwise. Each row represents a coefficient from a different regression. All regressions are instrumental variable estimations. Regressions also include full set of federal state, birth cohort and survey year fixed effects, state specific cohort trends, as well as male dummy and quadratic in age. The number of observations is in square brackets. Robust standard errors are clustered at the state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.6: Robustness checks

	Sample I		*Sample II*	
	Schooling	F-stat.	Maternal schooling	F-stat.
Main IV estimates from Tables 1.3 and 1.4	–0.063**	110.3	–0.068**	77.0
	(0.031)		(0.032)	
	[13,198]		[12,170]	
Panel A: Specification and sample restriction issues				
1. Leaving out the first affected cohort	–0.068*	83.0	–0.112***	52.8
	(0.038)		(0.034)	
	[12,718]		[11,773]	
2. Region-specific year of birth effects†	–0.092**	34.3	–0.101***	62.6
	(0.044)		(0.033)	
	[13,198]		[12,170]	
3. Interview mode	–0.064**	107.8	–0.070**	76.7
	(0.032)		(0.032)	
	[13,198]		[12,170]	
Bandwidth around the reform date:				
4. Six years	–0.087**	75.5	–0.028	31.2
	(0.038)		(0.044)	
	[5,711]		[4,930]	
5. Eight years	–0.085***	92.7	–0.034	69.9
	(0.031)		(0.036)	
	[7,562]		[6,260]	
Additional controls				
6. Political environment when 15	–0.067**	170.7	–0.054*	69.1
	(0.028)		(0.031)	
	[10,356]		[8,984]	
7. Economic environment, ages 12-18	–0.066**	96.0	–0.066**	39.9
	(0.029)		(0.032)	
	[8,063]		[6,979]	
8. Economic environment, survey year	–0.064**	108.2	–0.066**	77.2
	(0.031)		(0.032)	
	[13,198]		[12,170]	
Panel B: Alternative outcomes				
9. Ordered IV probit	–0.152**	110.9	–0.168**	76.5
	(0.077)		(0.068)	
	[13,198]		[12,137]	
10. Proportion of years being very concerned††	–0.043*	87.7	–0.054**	76.2
	(0.023)		(0.022)	
	[13,198]		[12,170]	
Panel C: Placebo test				
11. Placebo reform	0.202	2.5	0.230	0.5
	(0.205)		(0.509)	
	[13.198]		[12.170]	

Notes: The outcome is a dichotomous variable which takes on the value one if the individual is very concerned about immigration and zero otherwise. All regressions are instrumental variable estimations. Regressions also include a full set of federal state, birth cohort and survey year fixed effects, state specific cohort trends and male dummy. *Sample I* specifications also include a quadratic in age. †Regional split is as follows: Northern Germany: Hamburg, Bremen, Schleswig-Holstein and Lower Saxony; Central Gemany: North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Saarland; Southern Germany: Baden-Wuerttemberg and Bavaria. ††Includes full set of dummies of how many times the respondent answered to the outcome question and average age over the available panel years as a control for *Sample I*. The number of observations is in square brackets. Robust standard errors are clustered at the state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

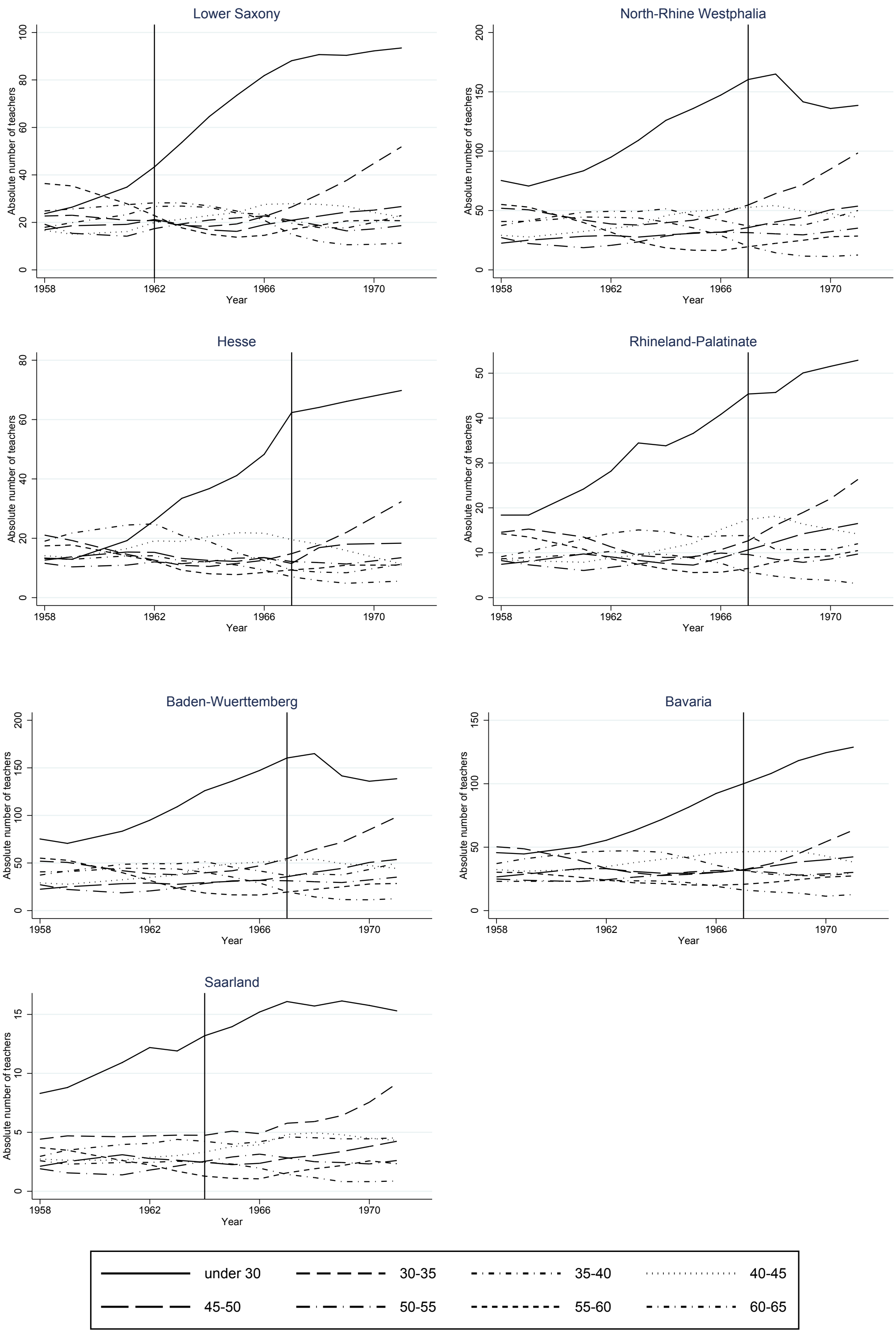
Table 1.7: The effect of schooling on potential channels

Sample I	Schooling	F-stat	Obs.
Panel A: Labor market outcomes			
1. Labor force status	0.017	110.2	13,198
	(0.028)		
2. Unemployed	–0.021	81.7	8,000
	(0.025)		
3. Ever unemployed	–0.036	89.9	12,088
	(0.029)		
4. White collar occupation	0.058	40.5	6,857
	(0.039)		
5. Civil servant	0.015	41.8	6,732
	(0.039)		
6. Unlimited labour contract	–0.053	39.4	7,003
	(0.036)		
7. Monthly labour earnings†	0.030	59.0	6,159
	(0.052)		
8. Household income††	0.044	97.4	11,926
	(0.037)		
Panel B: Concerns			
9. Very concerned about own job security	–0.036	49.9	6,978
	(0.025)		
10. Very concerned about own economic situation	–0.020	88.3	13,161
	(0.021)		
Panel C: Social capital			
11. Trust‡	0.176**	49.6	9,526
	(0.081)		

Notes: The outcomes are in row headings. Each row represents a coefficient from a different regression. All regressions are instrumental variable estimations. The regressions additionally control for a full set of federal state, birth cohort and survey year fixed effects, state specific cohort trends, as well as male dummy and quadratic in age. †Labour earnings refer to net salary in the previous month and are measured in logs. ††Monthly HH income is measured in logs. ‡Measures on trust are standardised with mean zero and standard deviation of one. Robust standard errors are clustered at the state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

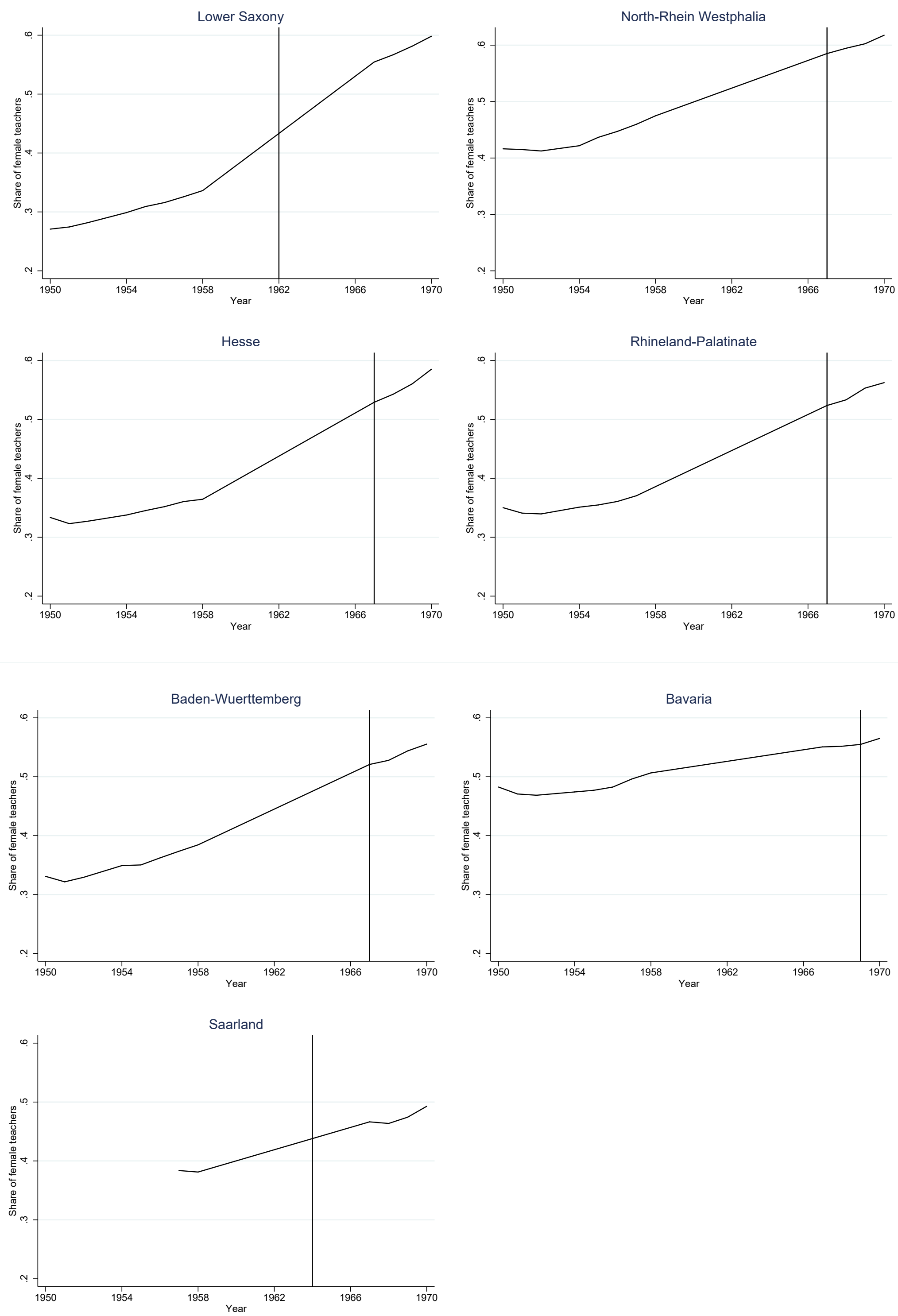
1.10 Appendix A: Additional Figures and Tables

Figure 1-4: Absolute number of teachers in basic track schools, in hundreds, by age groups and state



Source: Statistisches Bundesamt, Fachserie A: Bildung und Kultur; various series

Figure 1-5: Share of female teachers in basic track schools by state



Source: Statistisches Bundesamt, Fachserie A: Bildung und Kultur; various series

Table 1.8: Number of observations in each state by reform exposure

State	*Sample I*		*Sample II*	
	Not Affected	Affected	Not Affected	Affected
Hamburg	22	282	32	205
Schleswig-Holstein	176	422	183	351
Bremen	54	86	45	65
Lower Saxony	823	852	883	654
Saarland	26	7	3	30
Baden-Wuerttemberg	1,217	606	1,461	508
Hesse	840	408	883	258
North Rhine-Westphalia	2,598	1,182	2,468	801
Rhineland-Palatinate	678	331	677	211
Bavaria	1,954	634	1,993	459

Federal states are sorted in ascending order by year of reform.

Table 1.9: The effect of the reform on the track choice

Dependent variable	Reform	St. error	Obs.
Basic track	−0.011	0.017	13,198
Intermediate track	0.008	0.016	13,198
Academic track	0.003	0.013	13,198

Notes: The outcome is a dichotomous variable which takes on the value one if the individual attended the track mentioned, and zero otherwise. Each row represents a coefficient from a different regression. Regressions also include a full set of federal state, birth cohort and survey year fixed effects, state specific cohort trends, as well as a male dummy and quadratic in age. Robust standard errors are clustered at the federal state-year of birth level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.10: The curriculum for the 9th year of schooling in the various federal states

State	Summary of the content and goals	References
Hamburg	1. Competence in cultural technology 2. Overview of the political life 3. Vocational guidance	Nordrhein-Westfalen (1962), page 59
Schleswig-Holstein	**Historical-political topics for the 9th year** "The purpose of teaching history and politics education in the 9th year is directly related to current political, social, economic and cultural life, and to tasks which require a sense of responsibility and active participation." **Outline** 1. Help the workers: The social question 2. Our village (our city) as a political-economic community 3. Our home–Schleswig-Holstein–as a federal state 4. Democratic state under rule of law - dictatorship 5. Divided Germany - indivisible Germany 6. Towards European unity 7. Efforts to secure world peace in our time 8. Billions of people inhabit the globe: are they all brothers?	Engelhardt and Jahn (1964), page 188 Nordrhein-Westfalen (1962)
Bremen	"The 9th year of schooling in Bremen is dedicated in particular to a more in-depth general education of people, and primarily supports all measures serving this purpose."	Nordrhein-Westfalen (1962), page 60
Lower Saxony	"Political upbringing and education will help to fulfil the mandate of the legislature for schools to prepare the young people for life and work, and to make them committed citizens of a democratic and social legal state on the basis of Christianity, Western culture and German education.	Engelhardt and Jahn (1964), page 174

North Rhine-Westphalia	**Four educational tasks:** 1. Providing a more in-depth general education 2. Providing and reinforcing political education 3. Practising, acquiring and expanding basic skills and knowledge 4. Practical application	Nordrhein-Westfalen (1962), page 42
Hesse	**"The meaning and role of the 9th year** "To give young people a more in-depth general education and strengthen their character. Since the introduction of the 9th year, school can now provide a better political education than previously. Young people are becoming mature enough to formulate basic criticism, for "coexistence" and for fundamental insights: co-responsibility for the common good, preservation of human dignity, social justice and observance of the rules in the political struggle. The school environment and classroom style must correspond to these basic democratic values.	Engelhardt and Jahn (1964), page 7, 12
Baden-Wuerttemberg	**"From the education plan for the 9th school year"** "Even for a political consideration of political problems, young people have become more mature, so they can now gain a deeper historical awareness." ..."Bearing this in mind, recent history is to be examined once again in the context of political education." ..."Once treatment of the historical material has been completed to a certain extent by the end of the 8th school year, an integrated overview of history and social studies is now provided, aiming at democratic political education."	Engelhardt and Jahn (1964), page 165-166
Bavaria	**"The 9th year of schooling"** "This provides an in-depth insight into professional, economic and cultural life, strengthens young people's character, promotes active participation in the community and seeks to reinforce their values at a development stage when they are particularly impressionable."	Engelhardt and Jahn (1964), page 168

Notes: We were unable to find comparable information for the federal states Rhineland-Palatinate and Saarland.

Table 1.11: Timing of the reform and pre-reform state characteristics

	Dependent variable: Reform			
Student-teacher ratio basic track	−0.004			0.005
	(0.012)			(0.014)
Share migrants		−0.143		−0.140
		(0.122)		(0.119)
Share males		−0.007		−0.010
		(0.031)		(0.030)
GDP per capita		0.000		0.000
		(0.000)		(0.000)
Share unemployed		−0.049		−0.076
		(0.051)		(0.054)
Share CDU politicians			−0.005	−0.012
			(0.006)	(0.008)
Constant	0.080	0.494	0.136	0.870
	(0.403)	(1.582)	(0.205)	(1.447)
Observations	136	136	136	136
R-squared	0.541	0.570	0.544	0.582

Notes: The time period covers the year 1950 up to the reform year in each state. Due to unavailability of reliable data before 1950, all federal states, but Hamburg, are included. The dependent variable is a binary reform indicator that takes on the value zero throughout the pre-treatment period for each states, and switches to one once the compulsory schooling is introduced. Further treatment periods are discarded. The regressors are pre-reform state characteristics from the Federal Statistical Office and the Federal Employment Agency. All shares are measured in percent and missing values are replaced by the closest available value. The share of CDU politicians refers to the Landtag (state parliament). OLS regression with federal state and year fixed effects. Robust standard errors are clustered at the state-year level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.12: Interview modes by share of respondents

Interview mode	*Sample I*	*Sample II*
With interviewer's assistance	0.45	0.4
Oral interview	25.44	23.56
Self completed, with interviewer's help	16.28	22.36
Self completed, without interviewer's help	2.76	3.49
Oral and written	2.73	3.22
CAPI	47.60	38.96
Snail-mail	4.73	8.02
Phone interview	0.01	0.00
Proxy	0.01	0.00
Obs.	13,198	12,170

Source: SOEP, v30. Share of respondents is measured in percent.

Table 1.13: Dependent variable: immigration concerns

Explanatory variable		F-stat	Obs.
Panel A: Labor market outcomes			
1. Labor force status	0.009 (0.018)	107.1	13,198
2. Unemployment	–0.017 (0.029)	73.4	8,000
3. Ever unemployed	–0.015 (0.015)	85.8	12,088
4. White collar occupation	–0.123 (0.096)	27.7	6,857
5. Civil servant	–0.019 (0.046)	42.2	6,732
6. Unlimited labor contract	0.037* (0.019)	38.8	7,003
7. Monthly labor earnings †	–0.055 (0.044)	64.7	6,159
9. Household income††	–0.010 (0.040)	85.1	11,926
Panel B: Concerns			
9. Very concerned about own job security	0.194*** (0.038)	46.1	6,978
10. Very concerned about own economic situation	0.183*** (0.026)	87.0	13,161
Panel B: Social capital			
11. Trust‡	–0.082*** (0.015)	58.1	9,526

Notes: The outcome is a dichotomous variable which takes on the value one if the individual is very concerned about immigration and zero otherwise. All regressions are instrumental variable estimations. The regressions also control for full set of federal state, birth cohort and survey year fixed effects, state specific cohort trends, as well as male dummy and quadratic in age. †Monthly labour earnings refer to net salary in the previous month and are measured in logs. ††Monthly HH income is measured in logs. ‡Measures of trust are standardised with mean zero and standard deviation of one. Robust standard errors are clustered at the federal state-year of birth level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 2

Do Internships Pay Off? The Effects of Student Internships on Earnings[1]

Abstract

This paper studies the causal effect of student internship experience in firms on earnings later in life. We use mandatory firm internships at German universities as an instrument for doing a firm internship while attending university. Employing longitudinal data from graduate surveys, we find positive and significant earnings returns of about 6 percent in both OLS and IV regressions. The positive returns are particularly pronounced for individuals and areas of study that are characterized by a weak labor market orientation. The empirical findings show that graduates who completed a firm internship face a lower risk of unemployment during the first year of their careers, suggesting a smoother transition to the labor market.

[1]This chapter is joint work with Nils Saniter, Mathias Schumann and Thomas Sidler. I thank Fabian Gregor and the German Centre for Research on Higher Education and Science Studies (DZHW) for help and support with the data. Accepted for publication in the *Journal of Human Resources*.

2.1 Introduction

Internships have become a widespread phenomenon among university students in many countries throughout North America and Europe. Callanan and Benzing (2004), for example, argue that internships in the U.S. have become increasingly popular as a way to bridge the transition from education to work, with three out of four college students completing an internship in 2004, compared to fewer than 40 percent of students in 1980. In Germany, 55 percent of students who are currently enrolled in a university report having completed an internship during the past twelve months (Krawietz et al., 2006). By the time students finish their studies, nearly 80 percent report that they completed an internship while attending university (Sarcletti, 2009).

What motivates students to complete internships while enrolled at university? First and foremost, students expect internships to pay off after graduation when they enter the labor market. Indeed, when asked for their main motivation for undertaking an internship, most state the desire to get to know the work environment and gather practical work experience. Many also hope that an internship will help them to find employment later. The desire to earn money as an intern appears to be only a secondary motivator (Krawietz et al., 2006).

The surge in popularity of internships in higher education is not only a consequence of individual choices; it is also likely the result of universities emphasizing the importance of internships as part of the broader educational experience. Following the policy changes implemented as part of the Bologna Reform, graduates' employability has become a central objective of higher education across Europe (Teichler, 2011). Universities have been called upon to prepare their graduates better for the transition to work by focusing on competencies that are relevant to the job market. Internships are an effective means of building these competencies (Wolter and Banscherus, 2012; Teichler, 2011). As a consequence, many universities urge students to complete internships or even make internships an integral part of the curriculum.

Internships are believed to help students build work-relevant skills, gain specific knowledge of their future occupations, develop a clearer self-concept, and confirm or redirect individual career goals (Brooks et al., 1995). Most of the skills acquired during an internship are general and transferable (Busby, 2003). Students may further learn about their interests and preferences through internships and organize their remaining studies more efficiently by choosing courses satisfying those interests. These attributes may then translate into various favorable outcomes for the transition into the labor market and early career success, for example, shorter job search duration, lower probability of unemployment, more stable positions, better job match, and increased earnings. However, internships also produce costs due to the investment of time, effort, and sometimes even money. Interns have to accept educational opportunity costs, and might enter the labor market later than non-interns. Considering that most internships are poorly paid or not paid at all, it is not surprising that debate has arisen about the potential downside effects of internships. One argument is that firms exploit highly qualified students as cheap workers (Wolter and Banscherus, 2012). The overall effect of internships on individual labor market outcomes is unclear, and empirical research is needed to provide a basis for sound conclusions.

Theoretically, we anticipate student internships to have positive earnings returns. Human capital theory (Becker, 1993; Mincer, 1974) predicts that the additional knowledge, skills, and competencies accumulated as an intern result in higher pay if the time spent on the internship has a higher payoff for the specific career track than the time spent studying.[2] Signaling theories argue that employers' hiring decisions are made under uncertainty since the productivity of potential workers is unknown, and that job seekers may therefore use internships and positive references provided to them upon completion of the internship to signal high ability, which may result in improved job matching and higher earnings (Spence, 1973; Akerlof, 1970; Schnedler, 2004). Screening theory predicts

[2]We suspect that this is likely, given that most students do their internships between terms or semesters and therefore do not miss class.

that firms use such signals to more accurately assess workers' hidden productivity (Stiglitz, 1975). Social capital theory (Bourdieu, 1986; Coleman, 1988) also foresees positive labor market returns of internships because of the opportunity they provide to establish relationships with co-workers and potential employers. These social ties might, according to this line of thought, lead to better jobs after graduation (Granovetter, 1995).

We use longitudinal data from graduate surveys conducted by the German Centre for Research on Higher Education and Science Studies (DZHW) to empirically investigate the earning returns to internships. To account for the endogeneity of students' decisions to undertake an internship, we employ a two-stage least squares (2SLS) approach and instrument internship completion with mandatory internships. Exogenous variation comes from the introduction and abolition of mandatory internships across university and area of study level over time.[3] The first-stage regressions suggest that the presence of mandatory internships has a large and statistically significant impact on the likelihood of acquiring internship experience during university study. In fact, university students have a 56 percentage points (80 percent) higher likelihood of completing a firm internship during the course of their studies if the internship is mandatory. Internship experience increases earnings by around 6 percent, an effect that is precisely estimated in both OLS and IV regressions. Using an alternative instrument, that is the share of students reporting mandatory internships in a given university-study area-cohort cell, excluding the student's own report, we find that returns to internship are in the range of 14.6 percent. We interpret this as the upper bound of returns to internship experience. The positive returns are particularly pronounced for individuals

[3]We use the term 'area of study' to denote *Studienbereiche*, which are broadly defined areas rather than specific subjects or degree programs, for example Romance languages but not Italian, Spanish, or French. This is the term used by the German Federal Statistical Office and in our data. We use the term 'field of study' to denote a specific subject, such as Italian. The term 'study program' denotes a specific degree course, such as a master's program in Italian at a specific university.

and areas of study with a weak labor market orientation[4], and for humanities and social science graduates. Across other subgroups of the population, however, we do not detect heterogeneous effects.

We provide arguments and comprehensive evidence that mandatory internships are as good as randomly assigned, conditional on predetermined variables, such as area of study and university fixed effects. To support the credibility of the findings, several aspects are addressed: 1) measurement error in the instrument, 2) the risk of self-selection of students into areas of study with mandatory internships and 3) the impact of potential confounders, that is, simultaneity in the introduction of mandatory internships with other changes at the level of area of study or university. Importantly, detailed evidence from various student surveys shows that the requirement to complete an internship (or lack thereof) plays no role in students' choices of university or field of study (e.g., Hachmeister et al., 2007; Heine et al., 2009; Pryor et al., 2012). Furthermore, we analyze data from a representative internet survey conducted in Germany in 2016 that asked prospective university students and graduates in an open question about reasons for their study program choices and in another open question about reasons for their university choices. We additionally conducted a survey among first and third semester students at Universität Hamburg asking the same questions. We find no evidence that mandatory internships are relevant when choosing the university and study program, since no respondent answered that mandatory internships were important for their study choices.

The findings in this study have important policy implications. In particular, they suggest that combining university education with practical experience can enhance the job-market-relevant skills of students, especially for students in study programs with a weak labor market orientation, such as humanities.

The remainder of this paper is structured as follows: Section 2 gives a brief

[4]Labor market orientation is measured in two ways, first, by a graduate's self-assessment of how important labor market considerations were when choosing the study subject and second, by the occupational specificity of a graduate's area of study (see variables description in Section 2.3).

overview of the literature and the institutional setup of student internships in Germany. Section 2.3 describes the data, and Section 2.4 lays out the empirical strategy and discusses several aspects of identification. Section 2.5 presents the main results for the effects of internship experience on earnings and Section 2.6 inspects whether the effects differ for various subgroups of the population. Section 2.7 sheds light on labor market transitions, and Section 2.8 presents various robustness checks. Section 2.9 discusses limitations of the study and concludes.

2.2 Background

2.2.1 Related Literature

Despite the prevalence of student internships and their importance in allowing students to explore potential career paths, the empirical literature on causal effects of internship experience remains scant. Available studies draw conclusions based on opinion polls among interns about the perceived benefits of their work experiences (Beck and Halim, 2008; Cook et al., 2004; Shoenfelt et al., 2013; Krawietz et al., 2006). Another strand of literature compares individuals with and without internship experience, but does not account for potential self-selection into conducting an internship. Some studies also report positive correlations of internships with shorter job search duration (Gault et al., 2000), higher job stability (Richards, 1984), more and better-quality job offers (Taylor, 1988), a higher chance of choosing a career-oriented job (Callanan and Benzing, 2004), and wage increases (Gault et al., 2000; Reimer and Schröder, 2006; Sarcletti, 2009).

To our knowledge, the only studies that go beyond correlations are Nunley et al. (2016) and Klein and Weiss (2011). Nunley et al. (2016) conduct a résumé-audit study in the U.S. and randomly assign three-month internship experience to fictitious job seekers. They find that applicants with internship experience receive about 14 percent more interview requests than applicants without such experience. The effects are larger for non-business degree holders than for business degree holders. However, the authors are not able to study wage effects due to the experimental design. Besides, as the discussion in Heckman (1998) suggests, audit studies might find high returns to internship where there are none, and no returns where the returns do exist. Audit studies are based on the implicit assumption that the distribution of unobservables is the same in both groups after creating pairs that are "identical" in paper qualifications, hence comparing their outcomes by averaging over the outcomes at all firms for the same audit pair produces an estimate for the effect of internship. Heckman (1998) shows that if the variance of skills differs between the two groups, then which group gets the callback depends also on which group has a higher variance in relevant skills. Hence, conclusions drawn from audit studies depend on the distribution of unobserved characteristics for each group and the audit standardization level. Thus, the exploration of the earning effects of internships with observational data is important to advance our understanding.

Klein and Weiss (2011) use observational data from Germany for the graduation cohort 1997. They employ matching estimation methods and find no wage effects of mandatory internships among university graduates in Germany. Our study differs from Klein and Weiss (2011) in several aspects. Contrary to their paper, we follow the graduation cohorts and use two waves to increase the sample size. This allows us to exploit the variation in mandatory internships over time, making the identification more robust. Another important difference between our study and Klein and Weiss's study is the identified parameter: while our estimation approach identifies the local average treatment effect (LATE), the authors claim to estimate average treatment effect on the treated (ATT). The complier group in the present case are students who would have not done an internship had the internships not been mandatory. We believe that in this context the LATE is the more relevant policy parameter to identify. To explicitly compare our results to those of Klein and Weiss (2011) we replicate their study, additionally extending it to the cohorts we study in this paper. We report the results in Appendix C Table

2.18. Our replication shows that the differences in earning returns are rather driven by cohorts and not by the empirical approach.

2.2.2 Student Internships in Germany

There are two basic types of student internships: voluntary and mandatory. Whereas internships of the former type usually take place in firms or organizations outside the university setting, the latter may take place either in a firm or at the university. Students are generally free to do as many voluntary internships as they like during the course of their studies, but they have to fit these kinds of internships into their study schedule. As a result, students in Germany usually do voluntary internships during semester breaks.[5] In general, students do not earn university credit for doing voluntary internships.

In contrast, mandatory internships are mandated by study program regulations, and students earn credits for mandatory internships, that is, they have to complete an internship in order to graduate. The regulations state whether the internship has to be in a firm or at university. They further stipulate the industry and duration of the internship, and usually define learning targets. At the end of the internship, students must write a report about their tasks during the internship. To get an overview on the duration of mandatory internships, we scanned around 80 study regulations in Germany. We infer that most mandatory internships last 8-24 weeks (around two to six months). On average internships tend to be longer at universities of applied science.

Internships in firms may be either paid or unpaid. In our observation period, which was before the introduction of the general minimum wage in Germany in 2015, firms had no obligation to pay interns, irrespective of whether the internship was voluntary or mandatory. Even under the new minimum wage regulations,

[5] In Germany, an academic year consists of two semesters with a summer and winter break. The exact length of the breaks varies, but the summer break is usually 12 weeks and the winter break usually 8 weeks.

firms are not required to pay interns for voluntary internships shorter than three months or for mandatory internships. Nevertheless, firms and interns usually sign an internship contract stipulating the rights and obligations of both parties. Interns typically receive an internship certificate from firms including a performance evaluation after completion of the internship.

In this article, we are interested in the returns to internships outside of university including, for example, internships in firms and governmental and non-governmental organizations. We call internships outside of university 'firm internships' for brevity.

2.3 Data, Variables, and Descriptive Statistics

We use longitudinal data from the DZHW Graduate Panel, a survey of university graduates conducted by the German Centre for Research on Higher Education and Science Studies (DZHW). The DZHW Graduate Panel is described thoroughly in Rehn et al. (2011). Recent studies that have also used DZHW data are, for example, Parey and Waldinger (2011); Grave and Görlitz (2012); Freier et al. (2015). Each survey is a random sample of the student population of German universities. We employ information from three different cohorts comprising students who graduated in the years 2001, 2005, and 2009. For each cohort, an *initial survey* was conducted around one year after graduation. Around five to six years later, a *follow-up survey* was conducted. For the 2001 and 2005 cohorts, data are available for both waves, the initial and the follow-up survey. For the 2009 cohort, only the first wave is available. Figure 2-1 visualizes the timing of the data collection.

In the initial survey, students were asked whether they did a mandatory and/or voluntary internship during the course of their studies. In one question, each respondent was asked whether a firm internship was mandatory; respondents could tick the item "Yes, external internships (e.g., internship at a firm)". In another question later in the survey, each respondent was asked: "Which of the following means of further education and qualification, which go beyond your professional studies, did you use during your years of study?" Respondents could tick whether they did a voluntary firm internship. We use the first question to generate the instrument *Mandatory*, which is a dummy variable that takes the value one if a respondent stated that a firm internship was mandatory and zero otherwise. We use both questions to generate the endogenous variable *Internship*, which is a dummy variable that takes the value one if a respondent did a voluntary and/or mandatory firm internship and zero otherwise.

To check the reliability of self-reports we collect information on the existence of mandatory firm internships taken from study program regulations and—despite the strict data privacy rules—we were granted the permission to merge this information to a subsample of the DZHW survey data. We received a list of study programs with more than five respondents per study program from the DZHW for the graduate cohort of 2009. For this list, we collected as many study program regulations as possible and coded whether firm internships were mandatory. Afterwards, the DZHW merged the information given the regulations to the share of survey respondents stating that a firm internship was mandatory at the study program level.[6] Table 2.2 illustrates the results for study programs with and without mandatory firm internships (according to the study program regulations). The results show that the overlap of the survey responses with the study program regulations is substantial, supporting the accuracy of the survey responses.[7]

DZHW further collects information on details of graduates' university studies and their opinions about their studies. The surveys also include comprehensive demographic, socioeconomic, and educational information. In particular, the survey

[6]Note that merging the regulations to both the individual-level data and survey information other than the share of respondents stating that a firm internship was mandatory is prohibited due to German data protection.

[7]A perfect overlap is not expected, because conditioning on a respondent's university starting year was not possible. Given the substantial variation in the requirement to do firm internships over time, we expect that incorporating the university starting year would increase the overlap of the survey responses with the regulations.

collects various proxy variables for students' intelligence, ability, and labor market orientation, and information on parental backgrounds. The outcome variable, gross monthly earnings, is self-reported for the job at the time of the interview and measured in euros adjusted to 2005 prices. On average, we measure earnings around three years after graduating from university.

Throughout the analysis, we mainly work with a sample comprising all available waves of the three graduate cohorts as indicated by the shaded areas in Figure 2-1. This sample helps to increase the precision of the estimates, which will become particularly relevant when studying heterogeneous effects in Section 2.6.[8] In Section 2.8, we also distinguish between short-term (i.e., one year after university graduation) and medium-term (i.e., five to six years after graduation) labor market effects, using sub-samples of the available data.

A typical feature of some university degrees is that they imply an obligatory second phase of education. For example, prospective teachers in Germany take a first state exam upon completing their university studies, and then have to complete 1.5 years of classroom training before taking the state exam, which then qualifies them to work as a teacher. Similar obligatory second educational phases of varying duration exist, for example, for lawyers, clerics, and medical doctors in Germany. During this period, individuals are outside the regular labor market. For this reason, we exclude all individuals from our sample who finished university with a state exam (lawyers, clerics, pharmacists, teachers, and physicians) or reported having to complete an obligatory second phase of education. Furthermore, we exclude graduates who finished university with a bachelor's degree because of small sample size issues. Graduates with bachelor's degrees were interviewed only in 2009. Moreover, bachelor's degrees imply a shorter duration of study than other university degrees (Diplom, Magister, Master) and are less accepted by employers in Germany.[9] Finally, we only keep those observations that have non-missing values

[8]We borrow the idea of pooling the data from Parey and Waldinger (2011).

[9]In unreported regressions we estimate the returns to internships for a larger sample comprising all these groups, and the main results are similar.

for all relevant variables.[10] This results in a sample size of 13,976 graduates, with 19,736 person-wave observations. 6,790 graduates are observed in both the initial and follow-up survey.

Table 2.1 reports summary statistics for the overall sample (column 1) and differentiated by first- and second-wave observations (columns 2 and 3) and by internship experience in firms (columns 4 and 5). The numbers in column 1 in Table 2.1 show that 66 percent of graduates did a firm internship while attending university, 41 percent completed a voluntary firm internship and 48 percent a mandatory one.[11] The average year of birth is 1977, 53 percent are female, nearly one in three graduates completed an apprenticeship before starting at the university, and the average final high school grade is 2.2 (on a 1–5 grade scale with 1 signifying "excellent" and 5 "failing"). Furthermore, many students come from highly educated families, with 37 percent of mothers and 50 percent of fathers having graduated from an upper secondary school. Columns 4 and 5 indicate that students who did an internship express a stronger labor market orientation in their self-assessment when asked "To what extent did labor market considerations play a role when choosing your area of study?" (on a scale 1–5 with 1 signifying "not at all" and 5 "very much"). It is important to point out that labor market orientation refers to a point in time prior to entering university and can therefore be considered to be a predetermined variable. With respect to the outcome

[10] After excluding the initial groups we end up with 23,451 observations. We lose 2,460 observations because of missing values on the dependent variable, the instrument and the internship, ending up with 20,991 observations. The rest of the loss is due to missing values in parental education, high school graduation grade, and labor market orientation. In unreported regressions we replaced the missing values with the sample average and also control for it with dummy variables. The results are very similar and are available upon request.

[11] The corresponding figures in column 5 (0.63 and 0.70) do not add up to one, because many students did more than one firm internship during their studies. Among graduates who did a firm internship, about 37 percent did a mandatory but no voluntary internship, about 31 percent did a voluntary but no mandatory internship, and about 32 percent did both a mandatory and a voluntary internship during studies. Unfortunately, the DZHW data do not contain any further information on the number of internships. In Appendix B, we present descriptive evidence on the number of internships during studies from the Bavarian Graduate Panel, which is a survey similar to the DZHW survey.

variable—log monthly earnings—the unconditional means show that students who did an internship during the course of their studies have quite similar earnings to their fellow graduates.

Table 2.9 in Appendix A reports descriptive statistics for internship experience by field of study for three broad groups: (1) science, mathematics, engineering; (2) business and economics; (3) humanities and social sciences. It shows that more than 70 percent of graduates in humanities and social sciences acquire some internship experience while studying, followed by 68 percent of business and economics graduates. Figure 2-3 displays how distributed internship experience is across the universities in the sample. Similar to Table 2.1, we distinguish between internship experience (panel A), voluntary internship experience (panel B), and mandatory internship experience (panel C).

2.4 Empirical Strategy

We start by estimating the following OLS regression.

$$
\begin{aligned}
Log(Earnings)_{i,c,u,a,t+} = \beta_0 + \beta_1 Internship_{i,c,u,a,t} + \beta_2 GradCohort_c \\
+ \beta_3 AREA_a + \beta_4 UNI_u + X_{i,c,u,a,t}\gamma + \epsilon_{i,c,u,a,t}, \quad (2.1)
\end{aligned}
$$

where $log(Earnings)$ is the natural logarithm of gross monthly earnings of individual i of graduation cohort c of university u and study area a at time $t+$, where $t+$ is a time point after graduation. The variable $Internship$ equals one if a student did a firm internship while attending university at time t, and zero otherwise. We further include a rich set of fixed effects: $GradCohort$ includes dummy variables for graduation cohort, $AREA$ is a 53×1 vector of area of study fixed effects and UNI is a 262×1 vector of university fixed effects to control for the potential concern that the quality and reputation of the university and/or the study program may be correlated with the availability of mandatory firm internships and graduates' labor market outcomes later in life.[12] The vector X contains individual set of controls, such as a dummy variable indicating gender (*Female*), 22×1 vector of birth year indicators, a survey wave dummy, degree type dummy and several predetermined variables that are likely good proxies for students' intelligence and ability, in particular we control for students' final high school grade (*high school grade*), whether they completed an apprenticeship before beginning university studies (*apprenticeship*), the self-reported influence of labor market aspects on their choice of what career and study subject (*labor market orientation*), as well as a full set of dummy variables for mother's and father's highest general educational degree (four groups each).[13]

However, there are several reasons why the estimated effect of returns to internships may be biased in the OLS regression. It is likely that the choice to do an internship correlates with unobserved factors such as motivation or ability, that are also correlated with earnings later in life. Hence we rely on a 2SLS framework. To estimate the effect of internship experience on earnings later in life, we instrument firm internship experience with the presence of mandatory firm internships. The second-stage equation is:

$$\begin{aligned} Log(Earnings)_{i,c,u,a,t+} = \beta_0 + \beta_1 \widehat{Internship}_{i,c,u,a,t} + \beta_2 GradCohort_c \\ + \beta_3 AREA_a + \beta_4 UNI_u + X_{i,c,u,a,t}\gamma + \epsilon_{i,c,u,a,t}, \quad (2.2) \end{aligned}$$

[12]Note that for $AREA$, the data only allow us to observe the areas of study, which are referred to as *Studienbereiche* in the nomenclature of the German Federal Statistical Office, but not the exact subject (Statistisches Bundesamt, 2014). For example, we can observe whether someone studied Romance languages, but not whether the subject was French, Italian, Spanish, or Portuguese.

[13]Mincer type wage equations typically control for age and age^2 to proxy work experience. Age variables have been omitted from the baseline specification because they are likely to be outcome variables themselves. This is because internship experience might delay labor market entry due to the extra time working rather than attending university. We experimented with the inclusion of age variables and found that this leaves our results unchanged.

and the first-stage equation is:

$$\begin{aligned} Internship_{i,c,u,a,t} = {} & \alpha_0 + \alpha_1 Mandatory_{i,c,u,a,t} + \alpha_2 GradCohort_c \\ & + \alpha_3 AREA_a + \alpha_4 UNI_u + X_{i,c,u,a,t}\gamma + \varepsilon_{i,c,u,a,t}, \end{aligned} \tag{2.3}$$

where in the first-stage equation (2.3), the dichotomous variable $Mandatory$ equals one if a firm internship was mandatory during the course of studies, and zero otherwise. In the second-stage equation (2.2) the variable $Internship$ is the prediction from the estimated first-stage equation. The rest of the variables are defined as in equation (2.1).

2.4.1 Measurement Error in the Instrument

In the main specification, the instrument relies on individual survey responses on the existence of mandatory firm internships rather than on information from study program regulations. This approach might create measurement error in the instrument. Collecting all study program regulations from all universities to supplement and verify the accuracy of the survey data is, however, infeasible for two reasons. First, the regulations are either not readily available or unavailable, especially for older graduate cohorts. Second, restrictive data privacy rules in Germany generally prohibit the access to sensitive information such as university names, implying that the DZHW data do not comprise university names. We tackle the question of how well the survey responses match the study program regulations in two ways. First, we use alternative definitions of the instrumental variable in the spirit of Ashenfelter and Krueger (1994) to provide a range for internship effects. These instruments exploit variation in the requirement to do a firm internship at the department level over time, which is a major source of variation in our data. We construct departments by creating cells for each unique combination of a respondent's university and area of study. We further calculate the proportion of students reporting a mandatory firm internship by department and graduation cohort in a leave-one-out fashion. The second stage regression is identical to equation 2.2. The first regression then takes the following form:

$$Internship_{i,c,u,a,t} = \alpha_0 + \alpha_1 \frac{1}{k-1} \sum_{j=1, j \neq i}^{k} Mandatory_{j,c,u,a,t} + \alpha_2 GradCohort_c + \alpha_3 AREA_a + \alpha_4 UNI_u + X_{i,c,u,a,t}\gamma + \varepsilon_{i,c,u,a,t}, \tag{2.4}$$

where the second term in equation (2.4) is the ratio of people reporting the existence of a mandatory internship of graduation cohort c, university u, area a at time t, excluding the individual herself. We report the results from the above equations alongside to estimation results from equations (2.2) and (2.3).

Finally, we conducted our own additional short survey among students enrolled at Universität Hamburg. We asked first and third semester bachelor students the same question on the existence of mandatory firm internships as in the DZHW survey. Because we know whether firm internships are mandatory from the study program regulations, we can evaluate the accuracy of the question asked in the DZHW survey. Note that this test is rather strict, because we survey first and third semester students in contrast to graduates as in the DZHW survey—especially first semester students might not yet be aware of the obligation to do a (firm) internship. The results show that the survey responses are reliable. Out of 282 students, about 92 answered the question correctly. This result therefore further supports the accuracy of the survey responses.

2.4.2 Self-Selection into Study Programs with Mandatory Internships

Our identification approach crucially hinges on the assumption that individuals do not systematically select themselves into study programs with mandatory firm internships based on unobservable characteristics. Put differently, the instrument must provide variation that is exogenous given the control variables. This assumption would be violated if, for example, more ambitious students are more likely to choose study programs with mandatory firm internships, and if they are also more successful in the labor market later in life.[14]

The survey of the literature suggests that the quality and reputation of the study programs and universities (Hoyt and Brown, 1999; Parey and Waldinger, 2011) as well as proximity to one's nearest university (Spiess and Wrohlich, 2010) are likely to be the most important choice determinants. Several German national newspapers such as *Handelsblatt*, *Die Zeit*, and *Der Spiegel* regularly publish university rankings by subjects and institutions, and this information is widely circulated. Hachmeister and Hennings (2007) report that the majority of high school students in the final year of secondary school in Germany know and consult these rankings. However, none of these published rankings include information on internships. Moreover, gathering information from university websites on whether or not internships are mandatory is rather difficult, and unlike in the U.S., German universities do not distribute brochures or college catalogs to prospective students (Hoyt and Brown, 1999).

Table 2.10 in Appendix A summarizes studies asking students in Germany about factors that influence their study choices. The table provides an overview of the type of survey, sample size, the relevant question on study and/or university choice, key findings, and whether a question on internships is included in the survey. Most surveys elicit students' reasons for their university choice on a 5-point Likert scale (1 = "very important"; 5 = "not important at all"). For example, the representative surveys among first-year students conducted by the DZHW ask "How important are the following reasons for your choice of study?" on aspects such as reputation of the university, accessibility of the campus from home, quality of the academic program, etc. (Heine et al., 2005, 2009). Whether an internship is mandatory−or whether universities have good connections with

[14]Note that ambition would have to be an omitted variable that is not sufficiently captured by the predetermined observables such as high school grade, labor market orientation, and parents' educational background, all of which are included in the full model specification.

firms that enable students to find internships more easily during the course of their studies−was not among the items listed in the surveys. Indeed, none of the studies we have found on this topic for Germany lists mandatory internships as a relevant aspect of study choice (Hachmeister and Hennings, 2007; Hachmeister et al., 2007; Bartl and Korb, 2009; Institut für Marktforschung GmbH, 2014). Moreover, U.S. and Canadian student surveys indicate that internship availability does not play a role in students' choices (Hoyt and Brown, 1999; Canadian Undergraduate Survey Consortium, 2004; Pryor et al., 2012). These results suggests that educators, researchers, and empirical methodologists do not expect a mandatory internship to be a relevant aspect in students' choices of university or study program.

We further provide evidence from the German Internet Panel (GIP), wave 22. The GIP is a longitudinal panel survey conducted online on a bimonthly basis and is representative of the general population in Germany aged 16–75.[15] Survey participants who had already finished their studies were asked in an open question to name the three most important reasons for their university choice, and in another open question the three most important reasons for their study program choice.[16] Similarly, participants who had not (yet) started their studies were asked to name the three most important aspects that would be crucial for choosing a university and study program, respectively.[17] In total, 696 individuals responded to these questions. None of the respondents mentioned that mandatory internships offered by study programs were a factor in choosing the study program and university, and only two respondents (less than 0.3 percent) mentioned that

[15]The German Internet Panel is funded by the German Research Foundation through the Collaborative Research Center 884, "Political Economy of Reforms" (SFB 884). A survey description is available in Blom et al. (2015).

[16]The open questions are "If you attended university: Retrospectively, please name the three most important aspects that were decisive for your university choice(s)," and "If you studied: Retrospectively, please name the three most important aspects that were decisive for your study program choice."

[17]The open questions are "If you have not (yet) attended university: please name the three most important aspects that would be decisive for your university choice," and "If you have not (yet) studied: please name the three most important aspects that would be decisive for your study program choice."

internships completed before university helped them to choose the field of study.

In a similar vein, we conducted our own survey at the Universität Hamburg by handing out questionnaires to first and third semester students of various bachelor study programs. We surveyed students in their early bachelor studies to capture their motivations for choosing a certain study program and university as accurately as possible. The sample includes 282 students enrolled in programs both with and without mandatory firm internships according to the relevant study program regulation.[18] None of the survey participants indicated that the existence of mandatory (firm) internships was important for their choice of study program and university. Only one student indicated that a firm internship completed before studying played a role in study program choice.

Finally, we address this concern by using the DZHW data. We regress the instrument and its variations on a range of predetermined covariates. Table 2.3 shows that most coefficients are small in magnitude, statistically non-significant, and are not robust across different specifications in column (1)-(3), with the exception of having done an apprenticeship before starting university studies. The negative correlation with the instrument, however, is not surprising, as students who finished an apprenticeship before their studies are often exempt from doing a mandatory internship if their apprenticeships comprised field-relevant work experience. Finally, column (4) of Table 2.3 includes predicted earnings as a function of all observables as a dependent variable. The adjusted R-squared of the regression is virtually zero and the the estimate of the instrument itself is small in magnitude and statistically insignificant. We therefore believe that it is very unlikely that students choose their subjects and universities based on whether firm internships are mandatory. In summary, the evidence indicates that mandatory internships are not a decisive factor for students' university and study program choice.

[18] The sample includes four study programs with about 50–70 respondents each; while two programs require firm internships, two do not.

2.4.3 Impact of Potential Confounders

One threat to our identification strategy are changes at the university or department levels that affect earnings and coincide with the introduction or elimination of mandatory firm internships. For example, if the introduction of mandatory firm internships coincides with improvements in career counseling, we would overestimate the earning effects of internships. In order to assess the influence of such potential confounders, we examine twelve different indicators of study and university quality from DZHW that may affect earnings and thereby potentially bias the main results.

The twelve indicators cover the following four areas: 1) overall quality of education, 2) educational media and infrastructure, 3) training, and 4) career counseling. Respondents can rate items in category on a five-point scale, from "very bad" (1) to "very good" (5).[19] We test whether changes in the quality indicators across cohorts coincide with the introduction and elimination of mandatory firm internships by regressing the respective indicator on the set of controls specified in equation (2.3). The outcome takes the value one if respondents tick a four ("good") or five ("very good") on the respective measure, and zero otherwise.

Table 2.4 reports the estimates for the twelve different educational outcomes. Each coefficient and standard error in parentheses comes from a different regression. Positive coefficients imply that the presence of mandatory firm internships coincides with improvements in the quality indicators. 46 out of the 48 estimated coefficients coefficients in Table 2.4 are close to zero and statistically non-significant at conventional levels. The only estimates that are statistically significant are for the outcome variable *Up-to-date education* in columns (1) and (3).[20] We thus believe that concurrent changes in education quality are unlikely to drive

[19]Figure 2-4 in Appendix A displays the distribution of the twelve variables. The figure shows that there are considerable differences in how graduates evaluate the quality of their studies. For example, around 50 percent of the graduates rate the structure of the degree program and that the methods taught are up-to-date as good or very good (panel A). In contrast, fewer than 15 percent of graduates gave the same positive rating for career counseling (panel D).

[20]In unreported regressions, we also estimated logit models, which yield a very similar picture.

our estimates.

Our identification strategy also relies on the assumption that the departments' decision to introduce mandatory internships is unrelated to the underlying connections with the local labor market. If mandatory internships are linked to the departmental labor market attachment, these linkages would be part of the treatment, and hence our instrument. However, some of our results indicate that this is rather unlikely. In our heterogeneity analysis, we find that the effects are stronger in fields with weak labor market connections, such as humanities. This suggests that the effects are most likely to be driven by additional labor market experience. There is no reason to believe that the endogeneity would be concentrated specifically in these fields. Additionally, Table 2.4 shows that the quality indicators do not vary with the introduction of mandatory internships, which is also suggestive of no significant changes in the programs themselves.

Although we cannot definitively test this assumption, we present supplementary evidence in the Appendix, supporting a non-systematic connection between the mandatory internships and the local labor market. We test whether mandatory internships are correlated with the unemployment rate at the county level. To this end, we link the DZHW Graduate Panel to unemployment statistics from the German Federal Statistical office. Unemployment figures are calculated at the county level. In Panel A, we regress the self-reported mandatory internship dummy on the lags and leads of the unemployment rate relative to the graduation year. In Panel B, we run a similar specification at the department level, where we define a department as having a mandatory internship if 70 percent of the students of that cohort reported this to be the case. Table 2.12 presents the results. Across all specifications there is no evidence of a systematic association between the local employment rate and the existence of mandatory internships.[21]

[21] In unreported regressions we run the same specifications with logit and probit, and determine that there are mandatory internships at the department level if 60 percent of students report this to be the case. Additionally, we have regressed the mandatory internship dummy collected from the university websites on the general unemployment and youth unemployment rate at the level of the city where the university is located. The advantage of this approach is that the internship

2.5 Results

Table 2.5 presents the OLS and IV results. Each column shows the estimated coefficients and standard errors from a different regression. The first two columns present results for the OLS regressions, and columns 3-6 show the IV estimates. The standard errors are clustered at the department level.

All regressions in Table 2.5 show a positive and statistically significant relationship between firm internship experience while attending university and earnings later in life. The OLS coefficients for both specifications suggest that a student who gained labor market experience through a firm internship during the course of her studies has six percent higher earnings later in life.

The estimated coefficients are statistically significant at the 0.1 percent level. Importantly, the IV estimates also point to a causal positive and statistically significant relationship between firm internship experience and graduates' labor market earnings. In columns (3) and (4), where we instrument internship experience with self-reported mandatory internship existence, the estimates suggest a 6.5 percent return to internships. When we use the instrument with the leave-one-out ratio (IV_{II}), the coefficients indicate returns of around 15 percent. The difference in magnitude between the two IV estimates might be partly due to measurement error in the self-reported measure. Additionally, note that the second instrument assumes that there is no variation in mandatory internship requirement within a department for a given cohort, which is not always the case. Related to this, our second instrument might be capturing the internship returns to a different complier group. As the instrument is the share of students in a given department-cohort cell, it might be capturing the effect not only for individuals who did an internship because it was mandatory for them, but also for those who voluntarily choose to do so because of the high shares of internship takers. We report column (5) and (6) as an alternative specification to provide some range of the magnitude,

variable is not self-reported. However, we only have very few observations. The results are robust to all these manipulations and are available on request.

however we prefer the more conservative estimate, that is column (4) in Table 2.5 is our preferred specification.

Our identification strategy relies on the notion that a subgroup of the population will comply with the incentives set by the instruments. This implies that if an internship is mandatory, the population of compliers will complete a firm internship while at university, and in the absence of a mandatory internship, the population of compliers would not gain this practical experience. This concept is critical for our analysis, since the estimates can only be interpreted as causal effects for the population of compliers. Therefore, it is worth further examining the characteristics of the population of compliers. We follow the simple approach proposed by Angrist and Pischke (2009). For binary characteristics, the relative likelihood of a complier possessing the relevant characteristics is given by the ratio of the first-stage estimate for the specified group to the first-stage estimate of the entire sample. For instance, if we want to know the relative likelihood of the compliers being female or male, we conduct the first-stage estimates broken down by gender. In a second step, we can relate these group-specific point estimates to the first-stage estimate of the full sample to obtain the relative likelihood of the compliers being female or male, respectively. Table 2.13 presents the findings for eight different sub-populations for both instruments. Columns (2) and (4) provide the first-stage estimate for the relevant subgroup, while columns (3) and (5) give the relative likelihood of a complier belonging to the specified subgroup. The results in column (2) indicate that students complying with the mandatory internship are more likely to be male (with a relative likelihood of 1.09), and to have completed an apprenticeship prior to entering university (with a relative likelihood of 1.03). Compliers are also more likely to have a good or very good high school grade and to have parents who do not have an upper secondary school qualification (with a relative likelihood of 1.05). The results in column (5) of Table 2.13 show that the group of compliers for the second instrumental variable is slightly different. They indicate that complying students are less likely to have completed an apprenticeship, are less likely to have a good or very good high school degree, and are more likely to have parents with higher levels of education.

Overall, these findings suggest that there are some differences in the group of compliers between the two instruments. This might partly explain the difference in the magnitude of the effects we estimate. The difference in the instrumental variable estimates suggests that returns to internship are stronger for compliers who have not done an apprenticeship and had worse grades in high school.

First-stage results based on equation (2.3) are presented in Panel B of Table 2.5. The estimated coefficient for the instrumental variable *Mandatory internship* is always positive and precisely estimated at the 0.1 percent significance level. The estimates suggest that a mandatory firm internship increases the likelihood of firm internship experience by 56 percentage points. The corresponding F-statistics of about 2,542 point toward a strong first-stage relationship. Similarly, although the F-statistics for the second instrument is smaller than in column (4), it is nevertheless above the conventional threshold.

Table 2.5 also shows the estimated effects for other selected explanatory variables. Female graduates have around 17 percent lower earnings than male graduates. This estimated relationship is consistent with previous findings for Germany (Machin and Puhani, 2003; Leuze and Strauß, 2009). Moreover, the estimates for the variable *apprenticeship* reveals that graduates who completed an apprenticeship before starting their studies have around eight percent higher earnings.

In our main IV approach, the identifying strategy is based on instrumenting internship experience during studies by an indicator whether an internship was mandatory. In the interpretation of the Local Average Treatment Effect (LATE) framework, this suggests that IV estimates identify an effect for compliers, i.e., students who do an internship during studies because it becomes mandatory and would not have completed any internship during the course of their studies otherwise. The subpopulation for which the effect is identified might therefore be different compared to the overall student population (Imbens and Angrist, 1994).

Hence, it is important to point out that the similarity in the magnitude of the OLS and IV estimates might not necessarily be indicative of no or little selection in the OLS regressions, but the consequence of the different subpopulations for which the effect is identified. Noncompliance in the present context would imply students dropping out of university as a result of internships becoming mandatory. In Section 2.8, we present descriptive evidence from administrative data from the Federal Statistical Office suggesting that students are not more or less likely to drop out of university because an internship becomes mandatory.[22] Hence, it is likely that we are close to the case of a one-sided full compliance and the IV estimates might therefore be close to the Average Treatment on the Nontreated. Heckman et al. (2006) point out that this parameter answers an interesting policy question. In the present context, the parameter is informative about the earnings gains for students who would not have done an internship and who are selected into internship at random.

The positive LATE could be driven by various, not necessarily mutually exclusive, mechanisms. For example, internship experience may help develop and/or signal job-related skills (Akerlof, 1970; Spence, 1973; Mincer, 1974; Becker, 1993). The practical job experience in firms might create job networks and social capital theory also predicts positive returns to internship experience (Granovetter, 1995). It is important to point out that we cannot disentangle the extent to which the positive internship experience are driven by human capital theory, signaling theory, and social capital theory, because these theories produce very similar predictions. Moreover, returns might differ between voluntary and mandatory internships. For example, returns of voluntary internships could be higher if students who are interested and motivated learn more from the practical experience compared to those who are forced to do an internship, but are not motivated. On the other hand, it could be that a mandatory internship makes a significant difference in case students realize the (unexpected) benefits of the practical work experience.

[22]We cannot study the proportion of those not complying with a mandatory internship in the DZHW data, because we work with a survey of university graduates.

Is a six percent increase in earnings with internship experience a comparatively small or large effect? To answer this question, it is helpful to compare our results with the empirical literature on causal wage returns of education (Heckman et al., 2006). For the U.S., Angrist and Krueger (1991); Acemoglu and Angrist (2000) report causal wage returns to schooling of around six to ten percent, and Oreopoulos (2007) estimates returns of around 13 percent. For Germany, the returns to schooling estimates vary between zero and ten percent (Becker and Siebern-Thomas, 2007; Pischke and von Wachter, 2008; Saniter, 2012). Comparing our results with the literature, it therefore appears that the returns to internship experience are roughly comparable to the wage returns of one more year of schooling and therefore quite significant in size.[23]

2.6 Heterogeneous Effects

This section studies whether subgroups of the population benefit differently from firm internships. In addition to the results of separate models for each subgroup, we also report the relevant p-values from interacted models for both the OLS and IV approaches in Table 2.6.

Panel A reports the impact of internship experience in firms separately for men and women. Differences in returns to internship experience may show similarities to, for example, differences in college degree returns, which are higher for women than for men (Jacobson et al., 2005; Jepsen et al., 2014).

Panel B investigates whether effects vary by parental education. The sample is split into two groups by whether one of the parents has an upper secondary school degree. This split is based on the idea that students with highly educated parents

[23]Note, however, that the local average treatment effects are estimated for different groups. The literature on causal returns to schooling estimates returns for individuals with low levels of schooling who are forced to acquiring more education because of an increase in compulsory years of schooling. In this study, we estimate earnings returns of internship experience for university graduates and it is important to point out that different instruments are likely to define different parameters (Heckman et al., 2006).

could benefit from their parents' social networks, irrespective of their own labor market experience. Hence, a firm internship might be more rewarding for students without such intergenerational networks.

In panel C, separate effects are estimated for graduates by their final high school grade, since students with good and very good grades are likely to have other unobservable characteristics (e.g., high motivation, intelligence, social skills) that might make them benefit more from a firm internship than students with lower grades. Further, due to their abilities, they might be more likely to participate in a firm internship of high quality and prestige, an aspect that we cannot observe.

Panel D analyzes the heterogeneity of internship experience across students' labor market orientation. Students for whom labor market aspects played a critical role in their choice of what to study might be more ambitious and motivated during their internships than students with a weaker labor market orientation, potentially leading to higher returns. Alternatively, firm internships might be particularly beneficial for students who have not given much thought to labor market aspects: internship experience in firms might help them to gain a clearer self-concept and develop better career plans.

Panel E separately reports returns to internship experience for areas of study with a *strong* or *weak* labor market orientation. Following Sarcletti (2009), areas of study have a *strong* labor market orientation if they lead to a particular profession. Examples are medicine and architecture, because nearly all medical students become doctors and most architecture students later work as architects. In contrast, areas of study with a *weak* labor market orientation do not necessarily lead to a particular profession. These areas teach more general skills that qualify graduates for a wide range of jobs. Examples are history, philosophy, and languages.[24] Finally, panel F separately shows the impact of firm internships by field of study for three groups: (1) science, mathematics, engineering; (2) business and economics; (3) humanities and social sciences.

[24]See Table 2.11 in Appendix A for a complete classification of areas of study into weak and strong labor market orientation.

The estimates in panels A, B, and C in Table 2.6 do not point toward heterogeneous effects of internship experience in firms by gender, parental background, or high school performance. In contrast, the point estimates in panels D and E suggest that firm internships are particularly beneficial for students with lower levels of labor market orientation. For example, the IV estimates in panel D suggest returns of around 12 percent for students whose labor market aspects did not play an important role in their choice of what to study compared to only two percent for those who took labor market aspects strongly into consideration. The difference of ten percentage points is statistically significant at the five percent level, as indicated by the p-value of 0.017 from the interacted model. In line with this finding, the estimates in panel E also point toward higher returns of internship experience in firms for graduates in areas of study with a weak labor market orientation, with the difference being statistically significant at the ten percent level (p-value of 0.08 from the interacted model). The heterogeneous effects by field of study in panel F are also consistent with those in panels D and E. The estimates suggest that graduates in the humanities and social sciences (without economics) have higher firm internship earnings returns compared to those who studied science, mathematics, engineering, business, or economics.

We conclude that those who benefit most from internship experience in firms are individuals with a weaker labor market orientation and those who study subjects with a weaker labor market orientation. One explanation for this heterogeneity of effects is that firm internships help students to develop a better understanding of their future occupation and a clearer concept of their own preferences. Moreover, for graduates in subjects with a weak labor market orientation, firm internships can help to establish contacts with potential employers, which may facilitate the screening of candidates when the subject itself is not a strong signal.[25]

[25]In unreported regressions, we also distinguished between students who graduated from a university versus a university of applied sciences. Studies at universities of applied sciences are more practically oriented, and the effect of internship experience in firms might therefore differ by the type of university degree. The regression results did not point toward heterogeneous effects.

2.7 Transition to the Labor Market

This section examines how internship experience in firms affects the transition to the labor market, specifically during the first years after university graduation. We use calendar information in the surveys to construct binary activity indicators for every month during the first five years after graduation. Monthly information is available for employment, unemployment, and full-time employment. We use these indicators as additional outcome measures to study transitions to the labor market after graduating from university. Figure 2-2 graphically displays the estimated coefficients of internship experience in firms for these activities from OLS and IV regressions. The vertical bars represent the 95 percent confidence intervals. Panel A in Figure 2-2 displays the effects of internship experience in firms on the probability to be employed. While there are no statistically significant effects during the first two years, later years exhibit positive coefficients, though only significant at the five percent level during the third year. Panel B reports estimates on the likelihood to be unemployed. The graph reveals that internship experience in firms decreases the risk of being unemployed during the first year. However, in later years, this effect levels off to nearly zero and becomes statistically non-significant in most regressions. Panel C in Figure 2-2 shows the results for being in full-time employment. This indicator is only defined for employed individuals in the respective month. The graph shows a higher propensity to be in full-time employment in most months, with statistically significant point estimates mainly between 20 and 35 months after entering the labor market.

Given that firm internships reduce the risk of unemployment in the short term, the question is how much of the internship earnings return can be attributed to the reduction in the risk of unemployment? To analyze this question, we include the unemployment duration after graduation measured in months as a covariate in the full IV specification. The earnings effect of a firm internship declines from around six percent to 4.6 percent.[26] The decline in the point estimate is consistent with

[26]The internship effect remains statistically significant at the 5 percent level in the full IV

the negative wage effects found in the unemployment scarring literature, which finds that a one-month unemployment spell decreases wages, on average, by about 1 percent (e.g., Arulampalam, 2001; Gregg and Tominey, 2005; Gangl, 2006; Mroz and Savage, 2006). Further, the results in Figure 2-2 are also in line with recent work by Nunley et al. (2017) who find in a résumé audit study that internship experience obtained while studying mitigates the negative effect of underemployment on employment prospects of college graduates in the United States.

Overall, the findings in Figure 2-2 suggest that firm internships raise earnings by helping graduates to find their first job faster with a higher chance of being full-time employed. Moreover, including unemployment duration as a covariate in the regression suggests that firm internships work partly by reducing unemployment, but that the earnings effect is not entirely the result of increased job experience.

2.8 Robustness Checks

In this section, we present sensitivity checks. Table 2.7 reports the results of sensitivity analyses based on the full model specification similar to the regressions in Table 2.5, columns (2) and (4). First we present the estimates with an alternative definition of the instrumental variable, that is the threshold 70/30. The 70/30 threshold defines cells as having mandatory firm internships if more than 70 percent of all graduates report that a firm internship was mandatory, and zero if less than 30 percent report that its mandatory. Note that this approach involves some measurement error as we only observe departments and not students' actual study regulations, which would be more precise. The IV estimate in panel A is positive and significant, albeit it is nearly three times larger than the estimate from our preferred specification.

In panel B, we consider that certain departments might differ in educational quality, connections to firms, or degree of support provided to students in finding

specification; unemployment duration is statistically significant at the 0.1 percent level.

high-quality jobs. To control for these potential differences, panel B in Table 2.7 reports the estimates when controlling for a maximum set of 1,494 department fixed effects (i.e., dummy variables for unique combinations of university and area of study). These fixed effects are added to the full model specification, which already comprises area of study and university fixed effects. Hence, there might be the risk that this model is over-specified. It turns out that the coefficient for internship experience in firms decreases, suggesting positive returns of around four to five percent.[27]

In panel C of Table 2.7, we present estimated results for the models only with department fixed effects (excluding area of study and university fixed effects), therefore exploiting the variation within departments (i.e., the variation through timing) without the potential problem of over-fitting. Similar to the point estimates in panel B, the estimates suggest positive earnings returns of internship experience during the course of studies of four to five percent.

There might be differences in labor market returns for the same area of study across the applied science or full universities. For example, a degree in economics might differ in terms of quality or labor market returns between universities and universities of applied sciences. To address this concern, the regressions in panel D additionally include fixed effects for interactions between area of study and type of university. Reassuringly, the estimates do not change much.

There is also the risk that the returns on internship experience in firms are confounded by other forms of practical work experience. For example, 48 percent of graduates report paid employment during the course of their studies that was related to their degree. Moreover, the requirement to complete a firm internship might affect whether students seek other forms of work experience that might be substitutes or complements for firm internships. The regressions in panel E of Table 2.7 include a dummy variable for whether graduates worked during the

[27]Note that the model specification including department fixed effects is not our preferred one because statistical power declines, which would render the analysis of heterogeneous effects in Section 2.6 infeasible.

course of their studies. The point estimates for internship experience in firms remain largely unaffected, pointing towards positive internship returns of around 6 percent.

Sample attrition might be a problem, as only 34 percent of individuals participating in the initial survey were also interviewed in the follow-up survey. To address this concern, panel F restricts the analysis to earnings information from the initial survey, that is, earnings measured one year after graduation. Panel G restricts the analysis to earnings information from the follow-up survey, that is, earnings measured five to six years after graduation. The IV estimates in both panels point toward positive effects of internship experience in firms on earnings of around six percent. We therefore argue that the main findings are unlikely to be biased by selected sample attrition, and we note that the differences in earnings in the short and medium term are not very large (1.1-1.5 percentage points). This is conceivable if one factors in that a poor labor market start could result in lasting disadvantage for college graduates. For example, Franz et al. (1997) find that, in Germany, individuals experiencing unemployment directly after apprenticeship receive lower wages later in life, and a range of papers shows that labor market entry conditions matter for earning and employment prospects (Altonji et al., 2016; Lange et al., 2013; Raaum and Røed, 2006).

To explore whether the estimated returns to internship might be driven by outliers, we winsorize earnings at the one percent level. The point estimates in panel H of Table 2.7 remain very stable. In unreported regressions, we also winsorized the top and bottom 0.5 percent of log(earnings) and also estimated level models including zero earnings. The results are in line with the main findings and are available upon request.

Finally, we examine the potential problem of selected attrition. First, in unreported regressions, we estimated linear probability and probit models on graduates' likelihood of participating in the second wave. We found no empirical evidence of differences in attrition rates between those with and without an internship experience. Second, to investigate whether mandatory internships affect the probability to graduate, we obtain additional data from the Federal Statistical Office of Germany. The administrative data contains the number of students in incoming cohorts of 2003, 2004 and 2005 for 70 departments for which we also collect information from the university websites whether at the time internship was mandatory in the department. We also obtain data on the number of students graduating in 2007-2011 for the same departments. Unfortunately, we do not know precisely when the graduating students started their studies, hence in Table 2.8 we present three different approaches of matching the starting cohorts to graduation cohorts. We test whether the proportion of graduates, calculated as the number of graduates divided by the number of incoming students, is statistically different between departments with a mandatory internships and departments without one. Panel A of Table 2.8 presents the average proportion of graduates and the p-value of the difference for each of the three incoming cohorts matched with the number of graduates after ten semesters, which is the regular study duration. The proportion of graduates is similar between the two columns and the p-value of the difference is well above ten percent. Panel B repeats the same exercise, however matches each incoming cohort to the number of graduates after 12 semesters. Again the differences is statistically not different from each other. Finally, panel C calculates the proportion of graduates by aggregating over all cohorts. Namely, we divide the aggregate number of students who graduate in years 2007-2011 by the aggregate number of incoming students in 2003, 2004, and 2005. The mean differences are statistically not different from each other. We conclude that the presence of a mandatory internship does not change the graduation rate per se.

2.9 Conclusion

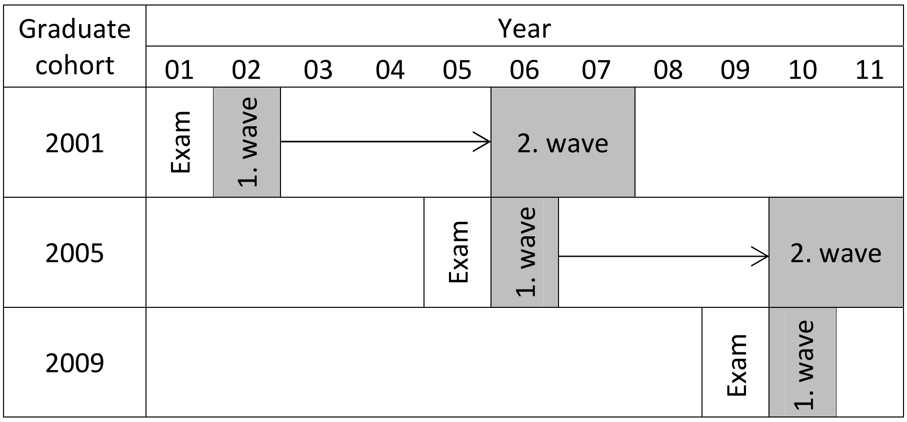
This study provides causal evidence of the effects of student internships in firms on earnings and transition to the labor market of university graduates. The estimates from instrumental variable regressions suggest that work experience gained through student internships in firms increases earnings by around 6 percent in the short and medium term. The empirical findings further suggest that graduates who completed a firm internship face a lower risk of unemployment during the first year of their careers. The positive returns are similar in magnitude for female and male graduates. There is also no empirical evidence of heterogeneous effects by students' socio-economic background and ability, proxied by their parents' educational attainment and students' average final secondary school grade, respectively. However, we do find significant differences in effects of internship experience with respect to the labor market orientation of students and the areas of study. The highest returns are estimated for a weak labor market orientation, and humanities and social sciences, which is in line with the notion of internships serving as a means of vocational exploration and screening.

The findings are confirmed by several robustness checks, and we provide comprehensive evidence showing that mandatory firm internships are likely to be exogenous. Due to data limitations, however, we cannot estimate longer-term effects as individuals' earnings are not observed 10–20 years after graduating from university. Further, we have no information on the size, sector, and reputation of the firm or institution at which the internship took place, and we can only present first suggestive evidence on the relevance of potential mechanisms. Hence, whether the returns vary by internship quality, firm characteristics, and on the extent and relevance of various possible mechanisms is left for future research. Despite these caveats, to the best of our knowledge, this is the first study aiming at estimating causal earnings returns to internship experience in firms. As such, this study complements the large empirical literature on the returns of schooling by estimating local average treatment effects of job experience among highly educated individuals.

The present findings are of interest to university students, policy makers, and educators alike. There is a growing debate in recent decades over the contradictory expectations placed on institutions of higher education: on the one hand, they are expected to incorporate labor market demands into their study curricula. On the other hand, they are expected to guarantee freedom and independence in academic research and teaching. Our study suggests that university education combined with practical learning—through firm internships—might be one way of bringing these two aspects [freedom of academic research "versus" demands of firms] together.

2.10 Figures and Tables

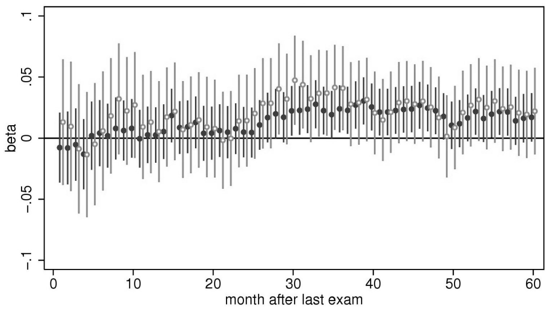
Figure 2-1: DZHW Panel Survey of Graduates



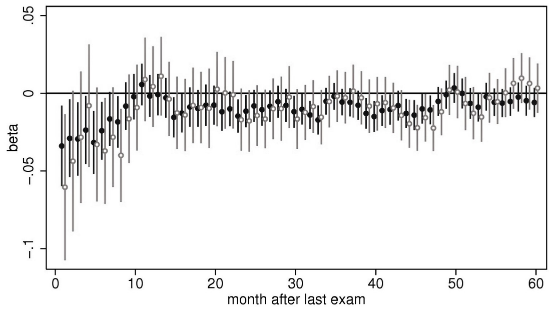
Note: Adopted from Rehn et al. (2011), p. 367. This study employs data from graduate surveys conducted by the DZHW. It includes random samples of university graduates who passed their last exam in 2001, 2005, or 2009. For the cohorts 2001 and 2005, we utilize an *initial survey* one year after graduation (first wave) and a *follow-up survey* about five to six years after graduation (second wave). For the cohort 2009, only the first wave is available. For the analysis, we use a pooled sample. It comprises all second-wave observations of the cohorts 2001 and 2005 and all first-wave observations of 2001, 2005, and 2009.

Figure 2-2: Transition variables over time

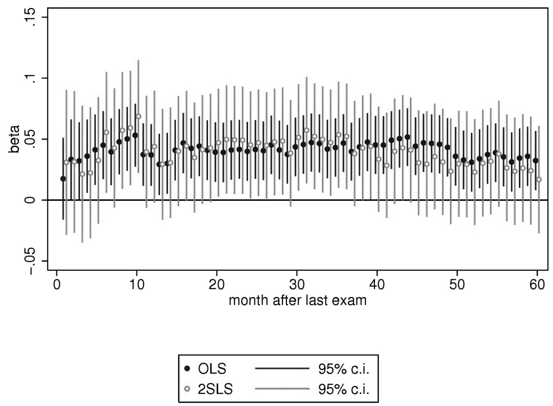
(a) Employed



(b) Unemployed



(c) Full-time employed



Note: Estimates are from OLS and 2SLS regressions (IV_I) for the effect of internship experience on binary variables indicating monthly status activity. Each circle represents the coefficient for one particular month. The vertical bars are the 95% confidence intervals. All models control for gender, year of birth FE, area of study FE, university FE, degree type FE, high school degree type mother and father FE, apprenticeship, high school grade, degree of labor market orientation and a dummy for the 2nd wave. *Data:* DZHW Graduate Panel.

Table 2.1: Sample means

	All	1 year after graduation	5–6 years after graduation	Internship: No	Internship: Yes
	(1)	(2)	(3)	(4)	(5)
Panel A. Explanatory variables					
Internship	0.66	0.66	0.67	0.00	1.00
Voluntary internship	0.41	0.41	0.41	0.00	0.63
Mandatory internship	0.48	0.48	0.48	0.00	0.70
Paid employment during studies	0.48	0.48	0.48	0.43	0.51
Year of birth	1977	1977	1976	1976	1977
Female	0.53	0.53	0.54	0.48	0.56
Apprenticeship	0.30	0.30	0.31	0.42	0.24
High school grade	2.23	2.24	2.22	2.26	2.22
Labor market orientation[a]	2.91	2.92	2.88	2.81	2.95
University of applied sciences	0.41	0.41	0.40	0.59	0.31
Mother has upper secondary school degree	0.37	0.38	0.36	0.32	0.40
— intermediate —	0.36	0.36	0.35	0.36	0.35
— lower —	0.26	0.25	0.28	0.31	0.24
— no —	0.01	0.01	0.01	0.01	0.01
Father has upper secondary school degree	0.50	0.51	0.50	0.44	0.54
— intermediate —	0.23	0.23	0.23	0.26	0.22
— lower —	0.25	0.25	0.27	0.29	0.23
— no —	0.01	0.01	0.01	0.01	0.01
Panel B. Outcome variable					
Log earnings	7.71	7.54	8.05	7.72	7.71
Employed	0.81	0.78	0.86	0.82	0.80
Unemployed	0.03	0.03	0.02	0.03	0.03
Full-time employed	0.70	0.65	0.79	0.70	0.69
Share of observations in 2nd wave	0.34	0.00	1.00	0.34	0.35
Number of individuals	13,976	12,946	6,790	4,720	9,256
Number of observations	19,736	12,946	6,790	6,631	13,105

Note: Column (1) presents variable means for the estimation sample according to Figure 2-1. Column (2) only includes observations from the first wave (1 year after graduation). Column (3) only includes observations from the second wave (5–6 years after graduation). Columns (4) and (5) divide the sample by treatment status. [a] The variable "labor market orientation" measures how important labor market aspects were with respect to study choice, measured on a five-point scale with 1 indicating "unimportant" and 5 "very important". *Data:* DZHW Graduate Panel.

Table 2.2: Survey responses and study program regulations

Share (%) of respondents stating that a firm internship is mandatory	Study program regulation: firm internship			
	Not mandatory (n=20)		Mandatory (n=17)	
	Share (%)	Cumulative (%)	Share (%)	Cumulative (%)
(1)	(2)	(3)	(4)	(5)
0	50.00	50.00	5.88	5.88
1–19	35.00	85.00	5.88	11.76
20–39	10.00	95.00	0.00	11.76
40–59	5.00	100.00	5.88	17.65
60–79	0.00	100.00	23.53	41.18
80–99	0.00	100.00	29.41	70.59
100	0.00	100.00	29.41	100.00

Note: The table combines responses on the existence of mandatory firm internships from the DZHW survey with information on the existence of mandatory firm internships gathered from study program regulations for a subsample of 37 study programs for the graduate cohort of 2009. The analysis includes study programs with more than 5 responses in the survey. Columns (2)–(5) show the relative and cumulative shares of survey respondents stating that a firm internship was mandatory per study program in percent (i.e., study program is the statistical unit). Columns (2) and (3) show the results for study programs where a firm internship was mandatory according to the study program regulation, and columns (4) and (5) show the results for study programs where it was not mandatory.

Table 2.3: Mandatory internships and individual characteristics

	Self-reported	Leave-one-out Ratio	Threshold 70/30	Predicted Earnings
	(1)	(2)	(3)	(4)
Female	0.002	0.010^{+}	0.011	
	(0.009)	(0.006)	(0.009)	
Share of observations in 2nd wave	0.000	0.001	0.002	
	(0.004)	(0.002)	(0.002)	
Apprenticeship	-0.060^{***}	-0.003	-0.017^{+}	
	(0.011)	(0.007)	(0.009)	
Mother has upper sec. high school degree	0.010	0.003	0.006	
	(0.009)	(0.004)	(0.006)	
Father has upper sec. high school degree	0.021^{**}	0.006	0.012^{*}	
	(0.008)	(0.004)	(0.006)	
High school grade	0.007	-0.009^{*}	-0.010^{+}	
	(0.007)	(0.004)	(0.006)	
Labor market orientation	0.011^{***}	0.002	0.002	
	(0.003)	(0.002)	(0.003)	
Mandatory Internship				0.005
				(0.020)
Cohort FE	Yes	Yes	Yes	
Birth year FE	Yes	Yes	Yes	
Area of study FE	Yes	Yes	Yes	
Degree type FE	Yes	Yes	Yes	
University FE	Yes	Yes	Yes	
Number of observations	19736	19271	13859	19736
Adj. R2	0.347	0.611	0.724	0.000

Note: The estimates in columns (1)-(3) are from regressions of the respective instrument on the covariates. In column (4) the outcome is the predicted earnings. Standard errors (in parentheses) are clustered at the department level. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. *Data:* DZHW Graduate Panel.

Table 2.4: Estimates of introducing mandatory internships on quality indicators

	Self-reported	Leave-one-out Ratio	Threshold 70/30
	(1)	(2)	(3)
Overall quality of education:			
Structure of the study program	−0.005	−0.018	−0.012
	(0.011)	(0.021)	(0.022)
State-of-the-art methods taught	0.001	0.013	−0.004
	(0.011)	(0.021)	(0.021)
Up-to-date education[a]	0.024*	0.032	0.036+
	(0.011)	(0.020)	(0.020)
Educational media and infrastructure:			
Availability of literature in the library	−0.009	0.009	0.000
	(0.011)	(0.021)	(0.022)
Access to IT services (internet, databases)	0.004	0.015	0.009
	(0.010)	(0.018)	(0.018)
Use of electronic communication devices	0.015	−0.024	0.014
	(0.012)	(0.022)	(0.022)
Training:			
Oral presentation training	0.004	0.021	0.017
	(0.011)	(0.023)	(0.022)
Writing skills training	0.000	−0.002	0.030
	(0.011)	(0.020)	(0.020)
Training in foreign languages[b]	0.005	0.013	0.002
	(0.009)	(0.018)	(0.017)
Career Counseling:			
Help in finding a job and starting a career	0.002	−0.021	−0.006
	(0.008)	(0.014)	(0.013)
Availability of career counseling	0.009	0.004	0.006
	(0.009)	(0.017)	(0.017)
Provision of career orientation events	−0.004	−0.003	−0.010
	(0.008)	(0.014)	(0.013)
Number of observations	18,220	17,789	12,782

Note: Estimates are from OLS regressions based on different definitions of the instrument. Standard errors (in parentheses) are clustered at the department level. All models control for gender, year of birth FE, area of study FE, university FE, degree type FE, high school degree type mother and father FE, apprenticeship, high school grade, degree of labor market orientation and a dummy for the 2nd wave. [a] The variable measures the actuality of education with respect to current job requirements. [b] The variable measures subject- or job-specific training in foreign languages. Note that departments in which 40–60% (or 30–70%) of graduates say that an internship was mandatory are excluded from the regressions, resulting in smaller sample sizes in columns (3) and (4). + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. *Data:* DZHW Graduate Panel.

Table 2.5: The effect of student internship experience on log earnings

	OLS		IV_I: Self-reported		IV_{II}: Leave-one-out	
	Base	Full	Base	Full	Base	Full
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS and second stage						
Internship	0.061***	0.061***	0.065***	0.065***	0.158*	0.146+
	(0.011)	(0.011)	(0.020)	(0.019)	(0.081)	(0.080)
Female	−0.165***	−0.170***	−0.165***	−0.170***	−0.169***	−0.175***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Apprenticeship		0.077***		0.077***		0.086***
		(0.013)		(0.013)		(0.016)
High school grade		−0.030***		−0.030***		−0.030***
		(0.009)		(0.009)		(0.009)
Labor market orientation		0.037***		0.037***		0.035***
		(0.004)		(0.004)		(0.005)
Panel B: First stage						
Mandatory internship			0.566***	0.560***	0.288***	0.284***
			(0.011)	(0.011)	(0.022)	(0.022)
F-statistic[a]			2,602	2,542	174	171
Adjusted R^2			0.460	0.466	0.244	0.257
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth year FE	Yes	Yes	Yes	Yes	Yes	Yes
Area of study FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree type FE	Yes	Yes	Yes	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes	Yes
Parental schooling FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.326	0.332	0.325	0.332	0.322	0.329
Number of observations	19,736	19,736	19,736	19,736	19,271	19,271

Note: In Panel A the dependent variable is $log(earnings)$. In Panel B the dependent variable is equal to one if a graduate completed an internship during the course of studies, and zero otherwise. Standard errors (in parentheses) are clustered at the department level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. *Data:* DZHW Graduate Panel. [a] Relates to the instrument variable "Mandatory internship" The dichotomous variable "Apprenticeship" is equal to one for graduates who completed apprenticeship training before entering university, and zero otherwise. The variable high school grade is measured on a 1–5 grade scale with 1 signifying "excellent" and 5 "failing". The variable "labor market orientation" measures how important labor market aspects were with respect to study choice, measured on a five-point scale with 1 indicating "unimportant" and 5 "very important".

Table 2.6: Heterogeneous effects

	OLS	IV_I	Number of observations
Panel A: Gender			
Women	0.075***	0.057+	10,523
	(0.017)	(0.031)	
Men	0.052***	0.079***	9,213
	(0.014)	(0.023)	
P-value of interaction (internship×women)	0.190	0.871	19,736
Panel B: Parental background			
Parents with 'high' levels of schooling	0.067***	0.071*	11,294
	(0.016)	(0.029)	
Parents with 'low' levels of schooling	0.045**	0.041	8,442
	(0.015)	(0.027)	
P-value of interaction (internship×highly educated parents)	0.330	0.253	19,736
Panel C: High school performance			
High school grade ≥ median	0.052***	0.043+	12,051
	(0.013)	(0.023)	
High school grade < median	0.070***	0.118**	7,685
	(0.019)	(0.038)	
P-value of interaction (internship×high grade)	0.100	0.068	19,736
Panel D: Labor market orientation of student			
LM orientation < median	0.070***	0.116***	7,351
	(0.018)	(0.032)	
LM orientation ≥ median	0.048***	0.018	12,385
	(0.014)	(0.025)	
P-value of interaction (internship×weak LM orientation)	0.050	0.017	19,736
Panel E: Labor market orientation of study subject [a]			
Strong LM orientation	0.049***	0.049**	14,743
	(0.011)	(0.018)	
Weak LM orientation	0.101**	0.139+	4,993
	(0.036)	(0.072)	
P-value of interaction (internship×weak LM orientation)	0.080	0.080	19,736
Panel F: Field of study subject			
Science, Mathematics, Engineering	0.050***	0.056**	10,125
	(0.014)	(0.020)	
Business and Economics	0.058**	0.040	3,921
	(0.021)	(0.047)	
Humanities and Social Sciences	0.085**	0.111*	5,445
	(0.030)	(0.056)	
P-value of interaction (internship×BE)	0.170	0.525	19,491[b]
P-value of interaction (internship×HSS)	0.086	0.086	

Note: All models control for gender, year of birth FE, area of study FE, university FE, degree type FE, high school degree type mother and father FE, apprenticeship, high school grade, degree of labor market orientation and a dummy for the 2nd wave. [a] See Table 2.11 in Appendix A for a classification of areas of studies into weak and strong labor market orientation. [b] 245 observations are outside of the study field categorization. Standard errors (in parentheses) are clustered at the department level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. *Data:* DZHW Graduate Panel.

Table 2.7: Robustness: Specification and sample selection

	OLS	IV_I	Number of observations
Panel A: 70/30 Instrument			
Internship	0.066***	0.164*	13,859
	(0.014)	(0.066)	
Panel B: Adding department fixed effects			
Internship	0.056***	0.043*	19,736
	(0.012)	(0.022)	
Panel C: Only department fixed effects			
Internship	0.056***	0.044*	19,736
	(0.012)	(0.022)	
Panel D: Area of study, university type, and interaction fixed effects			
Internship	0.057***	0.051+	19,736
	(0.011)	(0.028)	
Panel E: Employed during studies			
Internship	0.055***	0.062**	19,700
	(0.011)	(0.020)	
Panel F: Short-term earnings			
Internship	0.058***	0.069**	12,946
	(0.014)	(0.025)	
Panel G: Medium-term earnings			
Internship	0.063***	0.058*	6,790
	(0.014)	(0.024)	
Panel H: Winsorized earnings			
Internship	0.058***	0.062***	19,736
	(0.010)	(0.019)	

Note: All models control for gender, year of birth FE, area of study FE, university FE, degree type FE, high school degree type mother and father FE, apprenticeship, high school grade, degree of labor market orientation and a dummy for the 2nd wave. Exceptions: The regression in panel B omits area of study FE and university FE. Likewise, panel C omits area of study FE and the dummy indicating the university type due to the newly introduced interaction fixed effects between the two. Panel F uses earnings information only from the initial survey conducted around one year after graduation. Panel G uses earnings information only from the follow-up survey conducted around 5–6 years after graduation. Panel H winsorizes earnings at 1% level. Standard errors (in parentheses) are clustered at the department level. $^{+}$ p<0.10, * p<0.05, ** p<0.01, *** p<0.001. *Data:* DZHW Graduate Panel.

Table 2.8: The proportion of graduates by mandatory internship presence

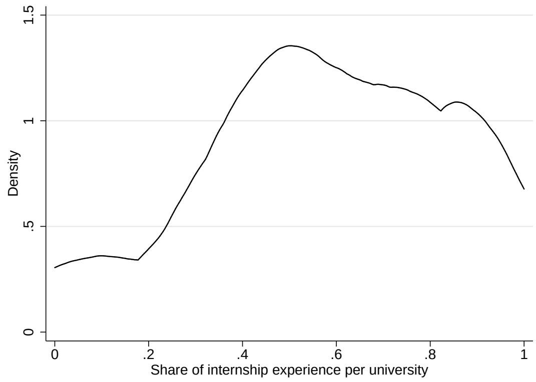
	Average proportion of graduates		
	Not mandatory	Mandatory	p-value of the difference
Panel A: After 10 semesters			
Incoming cohort 2003	0.536	0.648	0.215
Incoming cohort 2004	0.641	0.688	0.565
Incoming cohort 2005	1.038	0.748	0.375
Panel B: After 12 semesters			
Incoming cohort 2003	0.590	0.660	0.443
Incoming cohort 2004	0.564	0.602	0.614
Incoming cohort 2005	0.859	0.638	0.286
Panel C: Aggregated over all cohorts			
All cohorts	0.857	0.934	0.443
Number of departments	45	25	

Note: The table compares the proportion of graduates from departments with mandatory internships to the ones graduating from departments without mandatory internships. In Panel A the proportion of graduates is calculated by dividing the number of graduates after ten semesters from the respective starting date by the number of incoming students. In Panel B the enumerator is the number of graduates after 12 semesters. In Panel C the enumerator is the aggregate number of graduates in years 2007-2011, and the denominator is the aggregate number of incoming students in 2003, 2004, and 2005. *Source*: Federal Statistical Office.

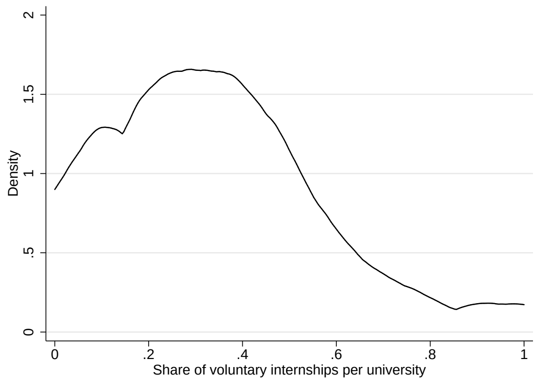
2.11 Appendix A: Additional Figures and Tables

Figure 2-3: Distribution of internship experience across universities

(a) Internships



(b) Voluntary internships



(c) Mandatory internships

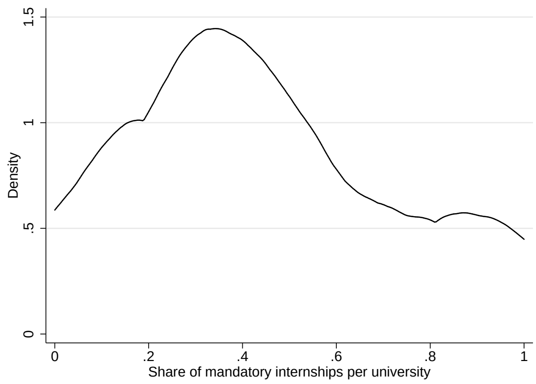
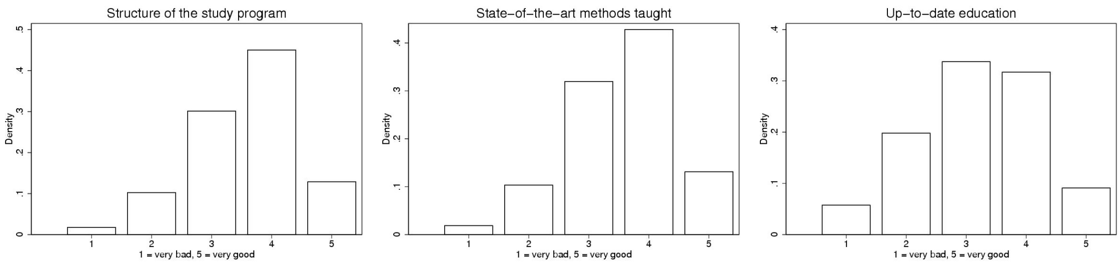
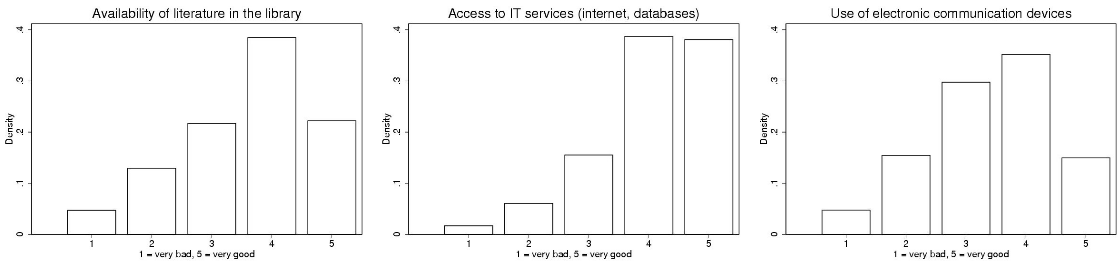


Figure 2-4: Students' evaluation of study related aspects

(a) Overall Quality of Education



(b) Educational Media and Infrastructure



(c) Training

(d) Career Counseling

Note: The corresponding questionnaire item reads "How do you evaluate the following aspects of your completed studies?" Respondents are then asked to answer on a scale from 1 ("very bad") to 5 ("very good"). Data is taken from the first wave ($n = 12,964$). *Data:* DZHW Graduate Panel.

Table 2.9: Internships by field of study subject

	Science, mathematics, engineering	Business and economics	Humanities and social sciences
Internship	0.62	0.68	0.74
Voluntary internship	0.33	0.51	0.51
Mandatory internship	0.50	0.40	0.50

Table 2.10: Overview of survey evidence on students' reasons for the choice of university and study program in Germany

Author(s)	Survey and sample	Question	Type of question	Main results	Question on internship?
Heine et al. (2005)	Representative survey of first-year students in Germany in the winter term 2004/05; 8,200 students.	How important are the following reasons for your choice of study?	5-point Likert scale (1 very important; 5 not important) on 20 different potential reasons/aspects with respect to the university and city.	Percent who answer that it is (very) important (1 & 2 on the scale): • Interest in content (91 percent) • Affinity/Ability (88 percent) • Many occupational choices later on (67 percent) • To be able to work independently (64 percent)	No
Hachmeister and Hennings (2007)	Survey of students in the middle of their studies.	How important were the following aspects for the choice of your university?	6-point Likert scale (1 very important; 6 not important at all).	Percent who answer that it is (very) important (1 & 2 on the scale): • To study preferred field of study (66 percent) • Good reputation of the university and professors (59 percent) • Proximity to home (58 percent) • Interesting city (51 percent)	No
Hachmeister and Hennings (2007); Hachmeister et al. (2007)	Survey among high school students in final grade; around 3,600 pupils.	How important were the following aspects for the choice of your university?	4-point Likert scale (1 applies very much; 4 does not apply at all).	Percent who answer that it is (very) important (1 & 2 on the scale): • Interest in content (100 percent) • Good facilities (90 percent) • Atmosphere in the city (89 percent) • Services for students (82 percent)	No
Bartl and Korb (2009)	Online survey among first-year students at the Martin-Luther University Halle-Wittenberg in the winter term 2008/09; around 800 students.	How important were the following reasons in your choosing this university?	5-point Likert scale (1 very important; 5 not important).	Average value of the 5-point scale: • Interest in content (1.79) • No need to pay tuition fees (1.96) • Good life conditions (2.23) • Proximity to home (2.27) • University has a good reputation (2.48)	No
Heine et al. (2009)	Representative survey among first-year students in Germany in the winter term 2007/08; 8,342 students.	How important are the following reasons for your choice of study?	5-point Likert scale (1 very important; 5 not important) on different reasons/aspects with respect to the university and city.	Percent who answer that it is (very) important (1 & 2 on the scale): • Interest in study program content (83 percent) • Proximity to home (66 percent) • Good reputation of the university (60 percent) • Good facilities (54 percent)	No
Institut für Markt-forschung GmbH (2014)	Representative online survey in 2013 among young people aged 16-24 who aim at studying at university; 500 individuals.	What is the most important aspect in the choice of the university?	5-point Likert scale (1 very important; 5 not important) on long list of aspects.	Percent who answer that it is very important (1 on the scale): • I do not want to be away too far from home (29 percent) • I will choose the university with the best study program (28 percent) • I want to live in a city that also has good recreational opportunities (23 percent) • I will choose a university with a good reputation (13 percent)	No

Table 2.11: Classification of areas of study into strong and weak labor market orientation

Strong LM orientation	Weak LM orientation
administrative studies	ancient/classic philology, modern Greek
agricultural sciences	area studies
architecture and interior design	arts, general art history
biology	catholic theology/religious education
chemical science	composition and design
civil engineering	cultural studies/cultural sciences
computer science	English studies, American studies
dentistry/dental medicine	extra-European linguistic and cultural studies
economics	film studies
electrical engineering	fine arts
engineering management	comparative literary and linguistic sciences
food and beverage technology	general cultural studies
forestry, forest and wood management	general economic and social science
general engineering	general linguistics and philology
geomatic/geospatial engineering	geography
geosciences (without geography)	German philology and studies
health-care science	history
human medicine	library science, documentation, communication
jurisprudence/law	music, musicology
landscape conservation, - architecture	education
mathematics, natural sciences	performing arts, theater studies
mechanical engineering, process engineering	philosophy
mining and metallurgy	political sciences
nautical science / navigation	protestant theology/religious education
pharmacy	psychology
physics, astronomy	Romance philology and studies
social pedagogy	Slavic, Baltic, Finno-Ugrian studies
spatial planning	social sciences
teletraffic engineering	special education
trophology, nutritional and domestic science	sport science
veterinary medicine	

Note: Based on Sarcletti (2009).

Table 2.12: Mandatory internships and local unemployment level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Panel A: Individual level regressions					
Lag 5	–0.006	–0.018	–0.029	–0.029	–0.024	–0.026	–0.023
	(0.007)	(0.016)	(0.018)	(0.019)	(0.021)	(0.022)	(0.021)
Lag 4		0.011	0.034	0.034	0.017	0.025	0.013
		(0.013)	(0.024)	(0.025)	(0.025)	(0.028)	(0.027)
Lag 3			–0.022	–0.022	–0.012	–0.019	–0.005
			(0.015)	(0.017)	(0.019)	(0.021)	(0.020)
Lag 2				0.000	0.016	0.021	0.024
				(0.015)	(0.033)	(0.034)	(0.034)
Lag 1					–0.019	–0.029	–0.035
					(0.026)	(0.032)	(0.032)
Lag 0						0.007	–0.017
						(0.012)	(0.019)
Lead 1							0.026
							(0.021)
Obs.	17,978	17,978	17,978	17,978	17,978	17,978	17,978
		Panel B: Department level regressions					
Lag 5	–0.006	–0.025	–0.009	–0.016	–0.008	–0.009	–0.008
	(0.013)	(0.019)	(0.022)	(0.027)	(0.029)	(0.029)	(0.029)
Lag 4		0.018	–0.018	–0.008	–0.022	–0.017	–0.020
		(0.016)	(0.030)	(0.036)	(0.041)	(0.043)	(0.044)
Lag 3			0.034+	0.023	0.034	0.030	0.032
			(0.019)	(0.028)	(0.032)	(0.038)	(0.039)
Lag 2				0.013	0.014	0.019	0.021
				(0.026)	(0.050)	(0.053)	(0.052)
Lag 1					–0.007	–0.015	–0.016
					(0.033)	(0.045)	(0.045)
Lag 0						0.005	–0.003
						(0.018)	(0.030)
Lead 1							0.009
							(0.027)
Obs.	1,294	1,294	1,294	1,294	1,294	1,294	1,294

Note: The outcome in Panel A is a dummy variable that takes on the value one for students reporting a mandatory internship, and zero otherwise. The outcome in Panel B is a dummy variable that takes on the value one if more than 70% of the students in a given cohort-department cell report a mandatory internship, and zero otherwise. Lags and leads are calculated in years with respect to each graduation cohort. The unemployment rate comes from the Federal Statistical Office and is calculated at the county level. All regressions include university and department fixed effects. Standard errors (in parentheses) are clustered at the department level. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. *Data:* DZHW Graduate Panel and Federal Statistical Office.

Table 2.13: Characteristics of compliers for the instruments IV_I and IV_{II}

Variable	IV_I: Mandatory (self-reported)		IV_{II}: Leave-one-out	
	First-stage estimate	Relative likelihood of compliers	First-stage estimate	Relative likelihood of compliers
(1)	(2)	(3)	(4)	(5)
Full sample	0.560		0.284	
Gender				
Female	0.510	0.911	0.258	0.908
Male	0.612	1.093	0.291	1.025
Apprenticeship				
Yes	0.576	1.029	0.254	0.894
No	0.552	0.986	0.297	1.046
High school grade				
Good or very good	0.587	1.048	0.258	0.908
Satisfactory or worse	0.514	0.918	0.319	1.123
Mother or father has upper secondary school degree				
Yes	0.539	0.963	0.309	1.088
No	0.587	1.048	0.252	0.887

Note: The table reports results for the characteristics of compliers for the two instruments IV_I (i.e., a dummy that equals one if a firm internship was mandatory during the course of studies, and zero otherwise) and IV_{II} (i.e., a dummy that is equal to one if 70 percent or more of the students at the department report that an internship was mandatory). Columns (2) and (4) report the first-stage estimates for subsamples as indicated by the dummy variables in column (1). Columns (3) and (4) report the relative probability that compliers have the characteristic indicated in column (1).

2.12 Appendix B: Internship Characteristics

The analysis thus far relies on data from the DZHW Graduate Panel. Although the DZHW data contain a plethora of information, one of the data's drawbacks is the lack of information on internship characteristics. We therefore supplement our analysis by using data from the Bavarian Graduate Panel (BAP), which is a survey of university graduates (similar to the DZHW Graduate Panel) in the German federal state of Bavaria.[28] Similar to our main analysis, we include survey information gathered one and five years after graduation and pool information from the graduation cohorts 2003/04 and 2005/06.

The BAP has several assets. First, it contains detailed information on internship experiences, which enables us to descriptively analyze internship characteristics. Second, the similar survey structures of the DZHW Graduate Panel and the BAP enable us to replicate our main results to a certain degree. Third, the BAP offers information on the field of study rather than area of study (e.g., one category for Italian instead of only one category for Romance languages).

Despite these advantages, the BAP has several shortcomings compared to the DZHW Graduate Panel for our analysis. First, the BAP only comprises information on graduates from Bavaria, which is only one of the 16 federal states in Germany. Because Bavaria has one of the strongest economies and labor markets in Germany, the BAP data might not be representative of Germany at large. Second, the BAP dataset is only about half the size of the DZHW Graduate Panel's—the smaller sample renders most of our previous instrumental variable analysis infeasible with respect to statistical power. Third, the wording of the questions on the existence of mandatory internships differs between the two surveys: while the DZHW Graduate Panel explicitly asks for mandatory internships *in firms*, the BAP does not. Fourth, the data do not include a university indicator; instead, we include indicators for the interaction of university type and study subject.

According to the BAP, students do, on average, 2.3 internships during the

[28] Detailed information on the BAP surveys is available at `http://www.bap.ihf.bayern.de`.

course of their university studies. Table 2.14 shows further descriptive statistics on the first and last internship during studies.[29] It shows that the duration of an internship is, on average, about four months, and that mandatory internships are longer than voluntary internships. Moreover, the first internship usually takes place in the third or fourth semester, and the last internship between the sixth and eighth semester.[30] The respondents were further asked to evaluate the first and last internship during their studies with respect to several dimensions. Notably, students evaluate mandatory internships to be more helpful than voluntary internships in giving them guidance on how to organize their studies. The table also indicates that students benefit more from mandatory internships in terms of job knowledge and skills.

Internships completed during the course of university study are a commonly used job search method. Figure 2-5 displays the share of graduates in the 2003/04 cohort who used a particular job search method to find their first job. Furthermore, the figure shows, conditional on having used a particular method, the share of graduates who assessed this particular method to be beneficial. The most common search methods are job advertisements, speculative applications, and contacts from jobs during studies. Contacts from internships during studies rank fifth—40 percent of graduates have used this job search channel. Moreover, more than 70 percent of graduates who used internships during their university studies as a job search method indicate having benefited from contacts acquired during that internship. Thus, a large share of university graduates make use of contacts from internships and find this channel to be a useful job search method.

As previously mentioned, participants in the BAP were asked whether internships during studies were mandatory; however, they were not asked whether internships *in firms* were mandatory. Nonetheless, Tables 2.15, 2.16, and 2.17 in Appendix A show that the OLS and IV point estimates are quite similar to those

[29] The information is available only for the cohort 2003/04.

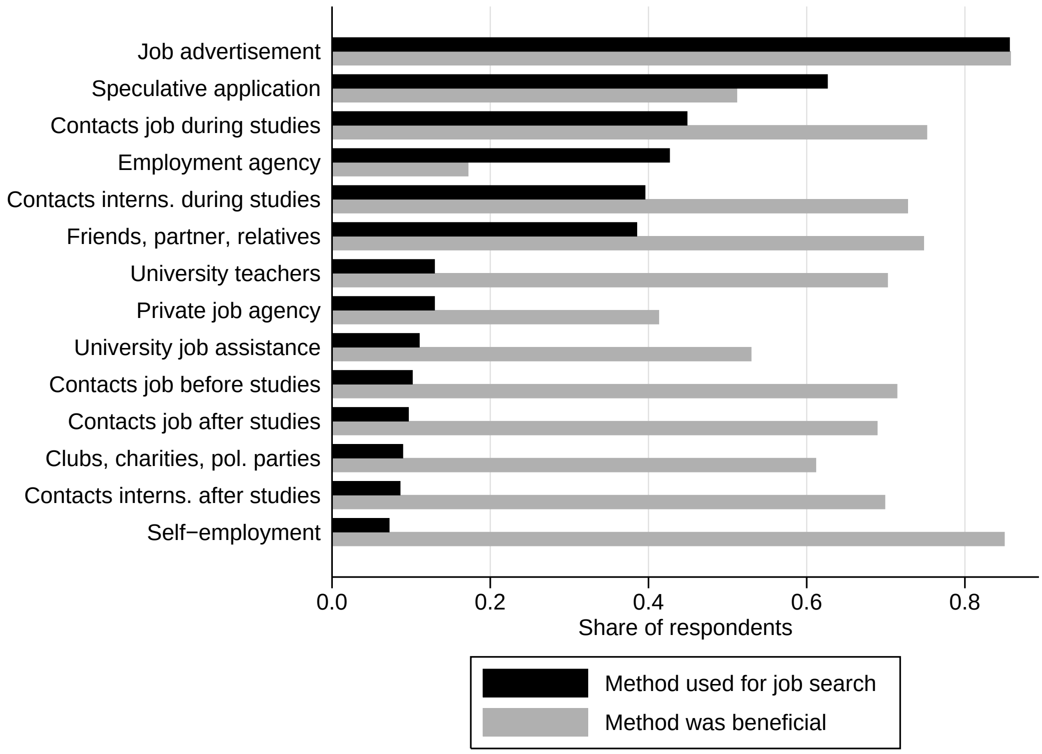
[30] Thus, the first internship usually takes place at the end of the second year of university and the last internship in the fourth year.

of the main analysis. In Table 2.15, the OLS coefficient of the internship indicator of the full model is 0.042 and statistically significant at the one percent level. The estimate implies that doing an internship while attending university increases earnings after graduation by 4.3 percent, thus the effect is about two percentage points lower than in the main analysis. The lower effect size is likely due to the lower importance of internships in the Bavarian labor market, which exhibits more or less full employment (in contrast to other federal states). Moreover, the coefficients of the covariates have the same sign as in the main analysis.

The IV results in columns (3) and (4) are based on an instrument that takes the value one if a student indicates having done at least one mandatory internship while attending university, and takes the value zero otherwise. The point estimates are somewhat lower and statistically non-significant. The non-significance is not surprising given the much smaller sample size. The low point estimates, however, warrant discussion. They are likely due to the kind of mandatory internships that the instrument captures. Instead of solely measuring internships in firms, the instrument is likely to also pick up internships at university to a certain degree. The smaller coefficient is nonetheless informative, because it points to a positive return of internships in general, but also highlights the additional value of completing an internship in a firm.

For the graduation cohort 2003/04, information is available on whether the intern's performance was assessed at the firm and/or university where the internship was completed. We use this information to redefine the instrument by assuming that interns who were assessed at a university but not at a firm did not complete a firm internship, that is, we assign internships that are unlikely to be a firm internship a zero in the instrument. Columns (5) and (6) show the results of this alternative instrument definition: the point estimates—although statistically non-significant—are larger and similar to the OLS results with a point estimate of 0.047 in the full model. Moreover, Table 2.16 shows that the first-stage relationship between the instrument and internship experience during studies is almost identical to the main analysis; and Table 2.17 indicates that internships have larger earnings returns in areas of study with lower labor market orientation. Overall, the results of the BAP confirm the results of the DZHW Graduate Panel.

Figure 2-5: Job search methods



Note: Respondents of the cohort 2003 who had already found their first job by the time of the first survey wave were asked which methods they used to find their first job ($n = 2,258$). Respondents could give multiple answers. The black bars indicate the share of respondents who used a particular search method. Conditional on having used a method, the gray bars show the share of respondents who found that this method was useful. *Data:* Bavarian Graduate Panel.

Table 2.14: Internship characteristics and perceived benefits

	Internship	Voluntary	Mandatory	p-value
	(1)	(2)	(3)	(4)
First internship during studies				
Duration (in months)	3.9	2.9	4.3	0.000
In/after which semester[a]	3.6	4.2	3.5	0.000
Internship benefits[b]				
Guidance for organizing studies	3.2	3.0	3.3	0.000
Better notion of job content	3.8	3.8	3.8	0.782
Job knowledge and skills	3.6	3.5	3.7	0.001
Soft skills	3.4	3.4	3.3	0.495
Making contacts for job entry	2.8	2.9	2.8	0.158
Number of observations	2,395	578	1,817	2,395
Last internship during studies				
Duration (in months)	4.3	3.4	4.7	0.000
In/after which semester[a]	6.7	7.7	6.2	0.000
Internship benefits[b]				
Guidance for organizing studies	3.4	3.1	3.6	0.000
Better notion of job content	4.3	4.3	4.3	0.889
Job knowledge and skills	4.1	4.0	4.1	0.155
Soft skills	3.9	3.9	3.9	0.832
Making contacts for job entry	3.4	3.4	3.4	0.807
Number of observations	1,718	576	1,142	1,718

Note: [a] In Germany, an academic year consists of two semesters. [b] The items are measured on a scale from 1 'not useful at all' to 5 'very useful'. Column 4 reports the p-values of two-sample t-tests comparing voluntary and mandatory internships. *Data:* Bavarian Graduate Panel.

Table 2.15: The effect of student internship experience on log earnings

	OLS		IV		IV firm	
	Base	Full	Base	Full	Base	Full
	(1)	(2)	(3)	(4)	(5)	(6)
Internship	0.040**	0.042**	0.020	0.028	0.033	0.047
	(0.014)	(0.014)	(0.020)	(0.020)	(0.044)	(0.044)
Female	−0.096***	−0.101***	−0.096***	−0.101***	−0.081***	−0.087***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.022)	(0.022)
University of applied sciences	−0.097	−0.113	−0.095	−0.111	−0.008	−0.098
	(0.080)	(0.081)	(0.080)	(0.080)	(0.155)	(0.156)
Apprenticeship		0.073***		0.073***		0.100***
		(0.012)		(0.012)		(0.023)
High school grade		−0.014		−0.014+		−0.002
		(0.008)		(0.008)		(0.016)
Labor market orientation		0.009*		0.009*		0.016*
		(0.004)		(0.004)		(0.007)
Cohort FE	Yes	Yes	Yes	Yes	No	No
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth year FE	Yes	Yes	Yes	Yes	Yes	Yes
Area of study FE	Yes	Yes	Yes	Yes	Yes	Yes
Degree type FE	Yes	Yes	Yes	Yes	Yes	Yes
Uni type*study area FE	Yes	Yes	Yes	Yes	Yes	Yes
Parental job qual. FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.285	0.287	0.284	0.287	0.220	0.225
Number of observations	11,603	11,603	11,603	11,603	4,494	4,494

Note: The dependent variable is log(earnings). Standard errors (in parentheses) are clustered at the individual level. $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data:* Bavarian Graduate Panel.

Table 2.16: First-stage results

	IV		IV firm	
	Base	Full	Base	Full
	(1)	(2)	(3)	(4)
Mandatory internship	0.592***	0.591***	0.525***	0.525***
	(0.012)	(0.011)	(0.021)	(0.021)
Female	0.026**	0.025**	0.016	0.015
	(0.008)	(0.008)	(0.014)	(0.014)
University of applied sciences	−0.063	−0.053	−0.018	−0.020
	(0.072)	(0.073)	(0.102)	(0.105)
Apprenticeship		0.008		0.000
		(0.008)		(0.013)
High school grade		−0.012*		−0.013
		(0.006)		(0.010)
Labor market orientation		0.002		0.002
		(0.003)		(0.004)
Cohort FE	Yes	Yes	No	No
Wave FE	Yes	Yes	Yes	Yes
Birth year FE	Yes	Yes	Yes	Yes
Area of study FE	Yes	Yes	Yes	Yes
Degree type FE	Yes	Yes	Yes	Yes
Uni type*study area FE	Yes	Yes	Yes	Yes
Parental job qual. FE	No	Yes	No	Yes
F-statistic[a]	2,664	2,659	618	617
Partial correlation coefficient	0.410	0.407	0.313	0.310
Adjusted R^2	0.536	0.537	0.478	0.479
Number of observations	11,603	11,603	4,494	4,494

Note: The dependent variable is equal to one if a graduate completed an internship during the course of studies, and zero otherwise. Standard errors (in parentheses) are clustered at the individual level. [a] Relates to the instrument variable "Mandatory internship". $^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. *Data:* Bavarian Graduate Panel.

Table 2.17: Heterogeneous effects

	OLS	IV	Number of observations
Panel A: Labor market orientation of study subject			
Strong LM orientation	0.033*	0.030	10,068
	(0.014)	(0.020)	
Weak LM orientation	0.106*	0.052	1,535
	(0.047)	(0.138)	
P-value of interaction	0.156	0.906	11,603
Panel B: Field of study [a]			
Science, mathematics, engineering	0.019	0.008	5,012
	(0.016)	(0.020)	
Business and economics	0.064*	0.056	4,098
	(0.029)	(0.046)	
Humanities and social sciences	0.082*	0.074	2,473
	(0.035)	(0.066)	
P-value of interaction (internship × BE)	0.196	0.509	11,583
P-value of interaction (internship × HSS)	0.144	0.377	

Note: All models control for gender, year of birth FE, area of study FE, Uni type-study area FE, degree type FE, university type, job qualification mother and father FE, apprenticeship, high school grade, degree of labor market orientation and a dummy for the 2nd wave. [a] See Table 2.11 in Appendix A for a classification of areas of studies into weak and strong labor market orientation. 20 observations are outside of the study field categorization. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$. *Data:* Bavarian Graduate Panel.

2.13 Appendix C: Replication of Klein and Weiss (2011)

In this section, we replicate the results from Klein and Weiss (2011) and extend their estimation strategy to additional waves. KW use data from the DZHW 1997 graduation cohort to asses the effect of mandatory internships on earnings five years after graduation. After sample restrictions similar to the ones in the present study, the authors end up with a sample of 1,971 observations. They focus on compulsory internships and use propensity score matching methods to reduce the potential bias naive OLS estimator might generate. They match on characteristics of the study program and on individual characteristics to estimate the propensity to be in a study program with a mandatory internship. The authors estimate an average treatment effect of -0.039 with a z-value of -1.58, implying negative but statistically insignificant returns to mandatory internships. In Table 2.18 in column (1) we replicate their findings. We report both unmatched OLS results as well as the estimated average treatment effect on the treated (ATT) after matching. Our sample is somewhat larger than KW's sample; nevertheless we get point estimates very close to theirs. In column (2), we again use the 1997 graduation cohort and propensity score matching methods similar to Klein and Weiss (2011), but this time we use as matching characteristics the variables we use in the present study. We show that the variable choice is not changing the results qualitatively. In column (3)-(5) we extend the sample to 2001, 2005, 2009 graduation cohorts, and in column (6) we pool these cohorts. The results clearly show that the experience of the 1997 graduation cohort is unique, as for all the other cohorts we get positive returns to internship, and in the full sample the effect is also statistically significant at five percent level.

Table 2.18: Mandatory Internships and Log Hourly Wages

Graduate Cohort	1997		2001	2005	2009	2001-2009
	Klein and Weiss					
	(1)	(2)	(3)	(4)	(5)	(6)
OLS (unmatched)	-0.047*	-0.040*	0.025	0.010	0.066*	0.027*
	(0.018)	(0.018)	(0.016)	(0.016)	(0.028)	(0.010)
ATT (prop. score matching)	-0.041	-0.034	0.040	0.045	0.080	0.054*
	(0.031)	(0.025)	(0.030)	(0.029)	(0.056)	(0.020)
Number of observations	2,089	2,046	7,229	8,826	2,639	19,415

Notes: The first column replicates the estimates of Klein and Weiss (2011), Table 3, column 3, page 982, using the same sample selection criteria and the same explanatory variables. KW report a point estimate of -0.039 [z-value$=-1.58$]. Column (2) uses the same sample selection criteria as in Klein and Weiss (2011) and similar explanatory variables as in our main specification (see Table 5, column 2), with the exception of university FE, because they are not included in the standard data format. Columns (3)-(6) report OLS and propensity score matching estimates for the different graduate cohorts, using our preferred specification (see Table 6, column 2). Estimates in columns (3)-(5) do not control for cohort and wave FE. Bootstrapped standard errors in parantheses. $^{+}$ $p<0.10$, * $p<0.05$, ** $p<0.01$.

Chapter 3

Low Emission Zones and Population Health[1]

Abstract

Air pollution has a detrimental impact on population health, but the effectiveness of policy measures targeting pollution is underexplored. I exploit the natural experiment generated by the staggered implementation of low emission zones in large cities across Germany to assess their impact on health. Using register data on outpatient and inpatient health care, I find that low emission zones reduce the number of patients with cardiovascular disease by two to three percent. This effect is particularly pronounced for those over the age of 65. The findings suggest that low emission zones can be an effective way to reduce air pollution and improve health.

[1]I would like to thank the Central Research Institute of Ambulatory Health Care (Zi), in particular Lars Kroll and Frank Ng, and Leibniz Institute for Economic Research (RWI) for their support in obtaining the data for this study. I am also grateful for the helpful discussions with Alexander Ahammer, Victoria Baranov, Miriam Beblo, Jakob Everding, Alexandra de Gendre, Luise Görges, Daniel Kamhöfer, Michael Kvasnicka, Jan Marcus, Eva Markowsky, Leslie Martin, Armando Meier, Nicolas Salamanca, Jonas Schreyögg, Hannes Schwandt, Thomas Siedler, David Slusky, and the seminar participants at the Johannes Kepler University of Linz. Accepted for publication in the *Journal of Health Economics*.

3.1 Introduction

Traffic accounts for more than one quarter of ambient air pollution in urban areas (Umweltbundesamt, 2018). Direct user charges, congestion pricing, license plate based restrictions, and low emission zones have become popular tools to reduce traffic-induced pollution (Davis, 2008; Wolff, 2014; Simeonova et al., 2018). Policies that aim to improve air quality primarily target population health: pollutants such as particulate matter, nitrogen oxides, and ozone can exacerbate already existing medical conditions, and in some cases even cause them. Despite the considerable attention paid to the traffic and pollution effects of air quality policies, their effectiveness in improving population health remains underexplored.

This paper studies the effects of one particular regulatory instrument, namely low emission zones (LEZs), on air pollution and cardiovascular health. LEZs are designated areas that restrict cars' access based on their emission class. Since 2007, multiple cities in Europe and specifically in Germany have gradually been introducing such zones. Across countries and cities, LEZs vary in their operating hours and the restrictiveness of their vehicle exclusion. LEZs in Germany are among the most restrictive, as they impose a 24/7 ban on all days, and the restrictions apply to all vehicles, with few exemptions. The rollout started in 2008 after it became clear that most major cities were not in compliance with the EU Air Quality Standards. In Germany, LEZs were introduced in three phases, with each phase excluding an additional emission class of vehicles. The rich temporal and spatial variation in the implementation of LEZs combined with their restrictiveness generates a compelling natural experiment to evaluate their impact.

To study the effect of LEZs on air pollution and cardiovascular health, I exploit the variation across space and time in the introduction of LEZs in German cities using a difference-in-differences design. I use air pollution measurements from the German Environment Agency covering the period 2004-2016. I focus on two criteria pollutants: particulate matter with an aerodynamic diameter of less than ten (PM_{10}) and nitrogen dioxide (NO_2), as vehicle exhaust is a dominant source of their emission.[2] I supplement the analysis with data from the Federal Motor Transport Authority for the years 2008-2017 to examine the impact of LEZs on the car fleet composition as a potential channel through which LEZs may reduce air pollution. To evaluate the impact on health, I use novel administrative outpatient care data from the universe of patients in the German statutory health insurance system during the period 2009-2017. I complement the outpatient data with evidence from a 70 percent random sample from German hospital diagnosis statistics for the years 2004-2014.

I find that LEZs reduce monthly PM_{10} concentrations by two to three percent. The reductions in monthly NO_2 concentrations are smaller in magnitude and often less precisely estimated. These effects corroborate the findings from previous evaluations of LEZs in terms of their effect on air pollution (Wolff, 2014; Malina and Scheffler, 2015; Gehrsitz, 2017). I further show that a likely channel through which LEZs affect air quality is by reducing the share of cars with the highest emissions in the treated cities.

By improving the air quality, LEZs also improve population health. Findings based on outpatient data suggest that the number of patients with cardiovascular disease decreases by 2-3 percent. Consistent with epidemiological studies, age and disease subgroup analysis shows that the effects are particularly strong for those above the age of 65 and for cerebrovascular diseases (7-12.6 percent). Event study plots show that changes in health outcomes occur roughly in parallel to changes in PM_{10}. Using complementary data from the Leibniz Institute for Economic Research (RWI), I further show that socioeconomic characteristics within LEZs do not change systematically over time compared to control areas, suggesting that estimates are not driven by compositional changes. Multiple robustness tests, such as the use of different control groups, sample restrictions, different definitions of outcome and treatment variables, and falsification checks suggest that

[2]Vehicle exhaust is a major contributor of carbon monoxide (CO) emissions as well. However, there are fewer CO monitors and CO concentrations are significantly below the limit values across the country.

the results are not sensitive to these alternative specifications. Findings from the hospital admission data also suggest a drop in the number of people admitted with cardiovascular disease, but with less precisely estimated point estimates.

Several epidemiological studies highlight the adverse association between air pollution, in particular PM_{10}, and cardiovascular health (Linn et al., 2000; Pope, 2000; Tsai et al., 2003; Brook et al., 2004; Wellenius et al., 2005; Franchini and Mannucci, 2007; Fiordelisi et al., 2017). This negative relationship can be observed even at levels below commonly applied concentrations thresholds. Notably, it is combustion-derived particles, and not atmospheric ones that carry hazardous compounds on their surface and are especially toxic (Mills et al., 2009). There are two main ways that particles trigger a health impact: they can cause a systemic inflammation or can translocate directly into the circulation. Regardless of the pathway, the inhaled particles trigger a range of biological responses, such as elevated heart rate, blood pressure, and heart variability, which cause the onset of cardiovascular symptoms.[3]

Cardiovascular disease generates the highest costs in the German health care system and is the leading cause of deaths worldwide.[4] The studies on the link between PM_{10} and cardiovascular health consistently show the strong association between PM_{10} and heart disease, and in recent studies, particularly between cerebrovascular disease (Cesaroni et al., 2012; Lee et al., 2014; Hong et al., 2002; Tsai et al., 2003; Wellenius et al., 2005; Lisabeth et al., 2008; Stafoggia et al., 2014). Cerebrovascular disease is especially interesting, as it is the second leading cause of mortality and the leading cause of morbidity (Leiva et al., 2013). Studies also show that the burden of cardiovascular risk from pollution exposure mainly falls on the elderly and individuals with pre-existing chronic medical conditions (Hong et al., 2002; Brook et al., 2010; Wellenius et al., 2005).

The adverse association between pollution and cardiovascular health can be

[3] For a detailed discussion see Mills et al. (2009).

[4] According to the German Federal Statistical Office, 13.7 percent (46 billion euros) of health care costs in 2017 were generated by cardiovascular disease (Statistisches Bundesamt, 2017).

observed both in the short term and in the long term. The evidence for the short-term link comes from episode and time series studies, showing that variations in air pollution are paralleled by changes in cardiovascular admissions and deaths, which are observed between a few hours and two days (Pope, 2000; Wellenius et al., 2005; Fiordelisi et al., 2017). Long-term exposure studies usually evaluate health outcomes across geographical area with different levels of average pollution over longer time periods. These studies evaluate the effects of persistent exposure as well as the cumulative effects of repeated variations in average pollution levels and find robust association between exposure and disease incidence (Pope, 2000; Stafoggia et al., 2014).

This paper contributes to the existing literature in several ways. First, despite the rich evidence available on the direct link between pollution and health, less is known about the general health effect of policy instruments targeting air pollution. This paper builds on this strand of literature by providing evidence on the effectiveness of low emission zones (LEZs) as a policy instrument. A concurrent working paper by Pestel and Wozny (2019) also analyzes the effects of LEZs in Germany. The authors find that LEZs reduce the number of circulatory and respiratory diagnoses among hospitals whose catchment areas have proportionally larger LEZ coverage. The present paper and the working paper by Pestel and Wozny (2019) differ in their empirical approaches, which is partly driven by the differences in the data used. Pestel and Wozny (2019) use data derived from hospital quality reports. Based on hospital location, the authors overlay its catchment area with LEZ coverage to compute the share of hospital catchment areas treated by the policy, since the quality reports do not provide information on patients' place of residence. Combining outpatient and hospital data allows me to show that it may be challenging to estimate precise treatment effects based on hospital data. I use individual-level hospital admission data, which contains information on patients' place of residence at the city level. I find that while effects point toward decreases in the number of hospital admissions, the coefficients are not precisely estimated.

Supplementing the analysis with outpatient data, which allows me to locate patients' place of residence at the postal code level, I show a significant drop in the number of patients with cardiovascular disease. Encouragingly, both papers find that LEZs reduce the number of cardiovascular disease. My findings on respiratory disease, however, suggest that the number of patients with this diagnosis has not changed significantly.

A second important contribution made by this paper is linked to the fact that most of the existing literature focuses on infant and children's respiratory health, while only a small number of papers also examine the effect on the elderly.[5] In an earlier evaluation of LEZs, Gehrsitz (2017) finds no effect on infant health. Two further studies use policy-induced variation to look at infant health by exploiting electronic toll collection in New Jersey and Pennsylvania (Currie and Walker, 2011) and childhood asthma by exploiting congestion pricing in Stockholm (Simeonova et al., 2018). This paper augments the literature by providing evidence for cardiovascular disease for the entire population and for the elderly. The elderly are of particular interest here because of their susceptibility to and the prevalence of circulatory system diseases.

Third, most studies evaluating the effect of air pollution on health use temporal variations in pollution levels over a short period of time (typically a day or a week), which allows them to estimate the impact of immediate exposure, but provides little information on long term exposure effects.[6] This paper adds to the small number of existing studies analyzing the effects of long-term marginal reductions in pollution in a setting with relatively low pollution levels (Gehrsitz, 2017; Simeonova et al., 2018; Alexander and Schwandt, 2019). The nature of the

[5]For papers on infant and child health, see Chay and Greenstone (2003); Neidell (2004); Currie and Neidell (2005); Currie et al. (2009); Neidell (2009); Lleras-Muney (2010); Beatty and Shimshack (2011); Coneus and Spiess (2012); Janke (2014); Sanders and Stoecker (2015); Simeonova et al. (2018); Alexander and Schwandt (2019), and for adult health, see Beatty and Shimshack (2011); Schlenker and Walker (2016); Deryugina et al. (2019); Isen et al. (2017); Gong et al. (2019); Anderson (2020).

[6]See, for example, Neidell (2009); Schlenker and Walker (2016); Deryugina et al. (2019).

present natural experiment and the availability of the rich and novel data over a long period enables me to evaluate health effects up to eight years after the policy was introduced. Furthermore, the national annual average level of PM_{10} in Germany was around 23 $\mu g/m^3$ in 2007 before the rollout of LEZs. Typical background concentrations of PM_{10} range between 20 and 50 $\mu g/m^3$ in developed countries and increase to between 100 and 250 $\mu g/m^3$ in developing countries (Mills et al., 2009).

The remainder of the paper is organized as follows: the next section provides the background on LEZs in Germany and Section 3.3 describes the data. Section 3.4 discusses the findings of this study for air pollution and car fleet adaptation. Section 3.5 presents the main results and Section 3.6 discusses the identification threats and the robustness checks. Finally, Section 3.7 concludes.

3.2 Low Emission Zones

The EU Clean Air Directives are among the strictest air quality standards worldwide. The first attempt to regulate the air quality in EU member states has been the directive 96/62/EC of the Council of the European Union, establishing the legal framework for ambient air quality regulation in all member states "to avoid, prevent or reduce harmful effects on human health and the environment as a whole". The directive defines how air quality should be assessed, specifies the pollutants for which alert thresholds and limit values should be introduced, and lists the measures member states should take to improve air quality. The daughter Directive 1999/30/EC establishes numerical limit values and alert thresholds for criteria pollutants. It divides the PM_{10} regulations into two phases: 2005-2009 and 2010 onwards with the aim to tighten PM_{10} regulations in 2010. Effective January 1 2005, the daily average of PM_{10} concentrations must not exceed 50 $\mu g/m^3$, and the yearly average should not exceed 40 $\mu g/m^3$. For the daily threshold 35 transgressions days are permitted. For NO_2, the daily average concentrations are limited to 200 $\mu g/m^3$, with 18 transgression days and the yearly average should not exceed 40 $\mu g/m^3$. The stricter regulations have not been phased-in, as most EU members struggled to meet the 2005 regulations.[7] Table 3.1 summarizes these regulations.

Between 2005 and 2007, 79 cities in Germany violated the daily threshold for PM_{10} and among them 12 violated the yearly threshold as well, according to estimations in Wolff and Perry (2010). The EU can impose significant financial penalties and even start infringement proceedings in case of non-compliance. Hence, the German government mandates that cities where even one pollution monitoring station violates the thresholds must develop a clean air action plan. Clean Air Action plans include four main elements: expanding public transportation, utilising ring roads, improving traffic flow and, most importantly, implementing LEZs (Wolff, 2014).[8] LEZs are designated areas that ban cars from accessing the zone based on their emission class.

LEZs are phased-in with increasingly stricter restrictions. Vehicle restrictions are motivated by EU-wide tailpipe emission standards that classify passenger vehicles in categories Euro 1 to Euro 6. Cars in the highest emission category (Euro 1) receive no windscreen badge. All other cars received color-coded badges. Phase 1 bans access only to vehicles with no badges. Subsequently, Phase 2 additionally bans vehicles with red badges (Euro 2), and Phase 3 further restricts access to vehicles with yellow badges (Euro 3). Cars in categories Euro 4 to Euro 6 receive green badges and are allowed in all stages.[9] The fine for violation is 40 Euro and a driving penalty point.[10] The policy mainly affects cars with diesel engines. Cars with petrol engine receive either no badge or a green badge. However, LEZs are

[7]Instead EU introduced new regulations for $PM_{2.5}$ in 2008 (directive 2008/50/EC).

[8]Wolff (2014) shows that the other elements of APs in general do not have a significant effect on PM_{10}.

[9]Table 3.11 in the Appendix presents details on tailpipe emission regulations. Table 3.12 presents some typical examples of cars in each Euro-class.

[10]The drivers lose their license after accumulating 18 points. However the fine for LEZ violation was replaced with 80 Euro and no penalty points in May 2014.

different from diesel driving bans, which are heavily discussed since 2018. Diesel bans restrict access to all diesel cars, with the exception of diesel Euro 6 category.[11]

Figure 3-1 plots the timing of these three phases. There is a vast temporal and spatial variation in the introduction of zones. The first zones were enacted in 2008. Among early introducers are Berlin, Hannover and Cologne, as well as cities in the Ruhr area, which in January 2012 united into a common LEZ for the Ruhr area, covering 869 km^2 and including 2.5 million residents. Not all LEZ-cities are non-attainment cities and vice versa: a number of non-attainment cities refrained from implementing LEZs.[12] Figure 3-2 illustrates the sample composition as well as spatial and temporal variation that I exploit in the empirical design. White perimeters represent the large cities which by 2017 have had no LEZ. The color-coded perimeters represent the large cities that introduced an LEZ in the observation period. Each color represents a group that enacted the zone in the given year. Towns and rural areas, shaded in gray, are excluded from the sample.

Table 3.13 in the Appendix presents the detailed list of the LEZ cities, with the enactment date, the areal coverage, and the attainment status of treated cities. Since each city is responsible for designing, enacting, and enforcing own LEZs, there is large variation in the proportional size of LEZs: the coverage ranges from as little as one percent of the city area to just below 100 percent covering the entire city. The perimeters are not randomly drawn–they mainly cover (potential) non-attainment areas of cities. Additionally, factors such as composition of the local fleet, urban layout, and social justice matter for designing the zones. A descriptive analysis suggests that in treated cities LEZs on average have slightly lower income compared to the rest of the city, have less cars per household and fewer residents

[11]Diesel bans are stricter than LEZs, but are rather unpopular in Germany. The biggest issue is with enforcing compliance. The problem is that currently all Euro 4, 5 and Euro 6 categories receive a green windshield badge, meaning that these cannot be differentiated either by the badge or the license plate. The only proof whether a vehicle is diesel is in the registration document. Hence, enforcement would require stopping vehicles for checks.

[12]Attainment and non-attainment in this context mean cities that comply with air pollution regulations and cities that do not.

over the age of 65. The most prominent difference between LEZs and the rest of the city is the population density, which is markedly higher inside the zones.[13]
Previous evaluations of LEZs in Germany all suggest that the zones reduce PM_{10} concentrations in treated cities. The first evidence comes from Wolff (2014), who analyses the short-term effects of LEZs until October 2008. He finds around a 7-9 percent drop in daily PM_{10} levels. Further evaluations support these initial findings, despite differences in time span and in composition of treatment and control units (Morfeld et al., 2014; Malina and Scheffler, 2015; Gehrsitz, 2017; Jiang et al., 2017) . In the most recent evaluation, Gehrsitz (2017) finds around a 2.5 percent reduction in daily PM_{10} levels on average, and around a 1.5 percent decrease in daily NO_2 levels. Gehrsitz (2017) additionally analyses potential effects on infant health, measured by birth weight, the incidence of low birth weight, and still birth. He finds a null effect.
It is worth noting that the diesel scandal happened in the middle of the observation period of this study. It is known now that Volkswagen and a few other car manufacturers have been manipulating the actual emissions of cars in laboratory settings through a software. The cars with clean diesel engine were introduced in 2008 and by 2017 there were 5.3 million such cars in Germany. Officially, these cars were mostly in Euro 5 and Euro 6 categories. The Federal Motor Transport Authority of Germany estimates that, on average, the cheating cars were emitting 59 percent more on the streets than their laboratory tests claimed. Alexander and Schwandt (2019) show the significant impact of the emission cheating on air pollution and health in the US. Due to this, it is likely that the pollution effects of LEZs shown in this paper are a lower bound.

[13]For details, see Table 3.14 in the Appendix.

3.3 Data

To study the effect of LEZs on air pollution, I combine data from multiple sources. I obtain daily PM_{10} and NO_2 measurements from the German Environment Agency (Umweltbundesamt) for the period 2004 to 2016. The sample consists of 264 PM_{10} monitoring stations and 261 NO_2 monitoring stations in 69 cities. Using the precise geographic coordinates of each station, I locate whether it is inside or outside an LEZ.[14] To account for weather-pollution interaction, I additionally collect weather data from the German Weather Service (Deutscher Wetterdienst, DWD). The data contain information on daily temperature, wind speed, precipitation, cloud cover, vapor, air pressure, and relative humidity. I match each air quality monitoring station to its geographically closest weather station and aggregate the daily measurements at the monitoring station-month level.

To study the health impact of LEZs, I draw on outpatient health data from the Central Research Institute of Ambulatory Health Care (Zi). These data comprise nationwide ambulatory care claims for patients of all statutory health insurance funds in Germany. The observations are annual and cover the period 2009-2017. This is a novel data source, which contains information on patients' postal code of residence. Because of data protection laws, it is not possible to obtain data on the postal code level directly. Instead, I obtain the number of patients with the diagnosis of interest, aggregated separately for control cities and by the areas inside and outside a low emission zone for treated cities. The diagnoses are coded according to the 10th revision of International Classification of Disease (ICD-10). Figure 3-3 illustrates the observation level with the example of Berlin. The gray shaded area covers the LEZ in Berlin. Hence, for Berlin, I obtain yearly aggregated data for the gray and white shaded areas separately. The advantage of this breakdown is that the treatment status can be defined very precisely. For cities that have no LEZ, I obtain the aggregated number of patients for the city

[14]The geographic coordinates are provided by the German Environment Agency. To classify the stations, I rely on open source polygons of low emission zones from OpenStreetMap.

as a whole.

My main outcome is the number of patients with cardiovascular disease (I00-I99). The data also allow to separately analyze heart disease (I20-I49) and cerebrovascular disease, including stroke (I60-I66), and to zoom into certain age groups. Of central interest are the elderly above the age of 65, however I also show results for the age groups 15-29 and 30-64. The data for younger patients is less reliable: the data confidentiality requires censoring all cells with less than 30 observations, which for patients under the age of 15 results in a large number of missing values.

I complement the main analysis with additional data from the hospital diagnosis statistics of the Federal Statistical Office. The inpatient administrative register comprises 70 percent random sample of all hospital admissions in 2004-2014, including emergency room admissions without overnight stay. The data provide information on date of admission and discharge, primary diagnosis, city of residence, as well as age and gender. However, it has two major caveats: first, identifying whether the patient lives inside or outside an LEZ is not possible. Hence, I define the treatment status by city of residence. Second, hospital admission is a severe outcome, particularly so in Germany, where hospitals are obliged to justify that an outpatient treatment is insufficient.

Panel A of Table 3.2 shows the means of PM_{10}, NO_2 and inpatient health outcomes in 2007 and reports the two sided p-value of the null hypothesis that the levels of these variables were the same in treated and untreated cities. Since LEZs are mainly measures employed in urban areas, I restrict the sample to cities with more than 100,000 residents to increase comparability between treated and untreated units. In this sample, the treated units are the cities that have introduced an LEZ during the observation period, and the control units are cities that have not introduced an LEZ by the end of the observation period.[15] I choose 2007, as it is the last year before the first rollout of LEZs. For proxying the baseline differences in health outcomes, I use inpatient data, as the outpatient data is available

[15]Note that the introduction of the zones continues and by 2019, many of the control units have enacted their LEZs as well.

only starting in 2009.

The annual average levels of PM_{10} and NO_2 in 2007 are, as expected, different between treated and untreated cities. LEZ cities have on average higher PM_{10} and NO_2 concentrations in 2007, and these differences are statistically significant. Note however, that even in non-attainment cities the annual concentrations of both pollutants are relatively low. Hence, results should be interpreted in the context of reductions against already low reference levels. In contrast to pollution, the prevalence of cardiovascular disease, measured as the number of hospital admissions per 10,000 inhabitants, appears to be statistically indistinguishable between treated and untreated cities.

It is straightforward to control for baseline differences, as long as the main identifying assumption holds. That is, that the trends between LEZ and non-LEZ cities do not differ systematically for reasons other than the implementation of the zones. To provide a suggestive evidence towards this assumption, Table 3.2 reports the *p*-values from a test of the null hypothesis that the year-on-year changes in all outcomes are different from zero. The results show that in the majority of cases the yearly changes in outcomes are not statistically different between treated and untreated cities before 2008. Event studies in Section 3.4 and 3.5 also show that pre-trends do not vary over time.

3.4 The Effect of LEZs on Air Pollution and Car Fleet

I first investigate how LEZs affect air pollution and car fleet composition, and then turn to health outcomes. Again, I restrict the sample to cities with more than 100,000 residents. To evaluate the effect of LEZs on pollution, I estimate the following equation:

$$y_{ict} = \alpha + \beta LEZ_{ict} + M_t + S_i + W_{ict} + \epsilon_{ict} \tag{3.1}$$

where y_{ict} is the monthly average concentration of PM_{10} or NO_2 at station i at time t in city c. LEZ_{ict} indicates whether a station i is located inside an LEZ at time t in city c. M_t and S_i are year-month and monitoring station fixed effects. Including station fixed effects accounts for time-invariant differences in the level of pollution between the stations (and, hence, also between the cities as the stations are nested in cities) and ensures that identification comes from within-station variation over time. Year-by-month fixed effects net out time shocks that commonly influence pollution in the cities. Furthermore, the vector W_{ict} includes a set of controls for weather, in particular, average temperature, its quadratic, maximum and minimum temperature, average humidity and its quadratic, an interaction term between average temperature and humidity, average air pressure, average precipitation and its quadratic, average wind speed, a dummy variable indicating rainfall, and an interaction term between average wind speed and average temperature. ϵ_{ict} is the error term. I cluster standard errors at the city level to allow for serial correlation within cities over time.

I estimate a variant of equation (3.1) excluding all pollution monitoring stations located in a treated city, but outside a low emission zone. These stations might be subject to negative spillover if drivers take longer tours to avoid LEZs, or positive pollution spillovers if the introduction of zones leads to a change in fleet composition or a drop in car usage in general.

Table 3.3 reports the results for PM_{10} in columns (1)-(3) and for NO_2 in columns (4)-(6). The results suggest that LEZs reduce monthly PM_{10} concentrations by 0.6-0.9 $\mu g/m^3$. The estimated coefficient is 0.9 $\mu g/m^3$ in column (3). This translates into a 3 percent decline in monthly PM_{10} relative to the average concentration levels in pre-LEZ period at treated stations. The findings for PM_{10}, being comparable in column (2) and (3), suggest negligible spillovers in either direction. The findings for NO_2 suggest a small reduction in its monthly concentration. After controlling for weather covariates, however, the effect becomes statistically insignificant.[16]

I focus on PM_{10} and not $PM_{2.5}$ as the number of $PM_{2.5}$ monitoring stations has been very small throughout the observation period, particularly for the period 2004-2010. Table 3.15 in the appendix presents the results for $PM_{2.5}$. The findings show that $PM_{2.5}$ decreased by around 0.47 μg/m^3.

To capture the dynamic effect of LEZs on pollution and gain more insights on the difference between treated and untreated cities in the pre-treatment period, I estimate an event-study model as follows:

$$y_{ict} = \alpha + \sum_{k=-5,k\neq-1}^{5} \beta_k LEZ_{ik} + M_t + S_i + W_{ict} + \epsilon_{ict} \tag{3.2}$$

where the dummy variables LEZ_{ik} indicate yearly lags and leads of up to five years before and after the introduction of LEZs. The reference category is the period -1, hence the effects are relative to the year immediately before the enactment. The rest of the controls are as specified in equation (3.1). Figure 3-4 presents the coefficients from the event study. The plots suggest that the common trend assumption is likely to hold, as all pre-LEZ coefficients are close to zero and statistically insignificant. The pre-LEZ coefficients are also jointly statistically insignificant. The respective p-values of an F-test for the joint hypothesis that all pre-LEZ coefficients are zero, are 0.74 and 0.63 for PM_{10} and NO_2. The event plots also suggest that the reductions in pollution levels are stronger from the third year onward, compared to the year just before the enactment of the zones.

These results are broadly in line with previous findings (Wolff, 2014; Gehrsitz, 2017). The lack of impact on NO_2 may reflect the slow improvements in nitrogen oxide emissions in real world conditions by diesel engines. Carslaw and Rhys-Tyler

[16]In unreported regressions I also run specifications where the air pollution is aggregated at the city level and separately by traffic stations and by background stations, as defined by the German Environment Agency. The results confirm a significant 1.2 $\mu g/m^3$ reduction in PM_{10} at the city level across traffic stations. The reductions across background stations and in NO_2 are small in magnitude and statistically not different from zero.

(2013) show that the reduction in NO_x emissions from all types of diesel vehicles over the past 15–20 years has been very small.

3.4.1 Change in Car Fleet Composition as a Potential Channel

There might be multiple channels through which LEZs affect the air quality. In the short term, LEZs are likely to reduce the absolute number of cars on the roads. Car owners with vehicles that do not comply with the regulations might switch to different means of transportation, such as public transport or bicycles. In the medium and long term, however, the absolute number of cars might increase again. Thereby, the composition of the fleet with respect to emission standards will likely be different, that is, the highest emitting cars will be replaced with cleaner ones.

This section studies the effect of LEZs on car composition. I draw on data from the Federal Motor Transport Authority (Kraftfahrtsbundesamt Flensburg). The data includes yearly observations on the total number of passenger vehicle registrations by city and emission class (Euro 1-Euro 6). I collect data for the period 2008-2017. The data is reported on January 1 of each year, hence the factual period the observations cover is 2007-2016. Note that the assigned color of the windshield stickers mainly depends on the Euro-class of the car, but can vary depending on the tax class of the car and the existence of a particle filter. Given this shortcoming of the data, it is likely that the share of no-, red-, and yellow-sticker cars is lower in practice, while the share of green-sticker cars is probably higher.

Following the implementation of LEZs, the composition of the car fleet has changed significantly. As Figure 3-11 in the Appendix shows, the share of Euro 1 cars has dropped from around 20 percent in 2007 to below two percent in 2017. In the meantime, the share of Euro 4 and higher class cars has increased from 40 percent to nearly 80 percent.[17]

To study the effect of LEZs on the composition of vehicle registration more formally, Table 3.4 reports the regression results following the same difference-in-differences empirical approach as in the previous section. The outcome variable is the share of the respective class of cars as specified in column headings in the total number of registrations in a city. The results support that LEZs reduce the share of Euro 1 cars and increase the share of Euro 4 cars in treated cities. The effect of LEZs on the share of Euro 4, 5 and 6 classes combined is positive, but statistically insignificant. This is conceivable, given that Euro 5 and Euro 6 categories, despite receiving a green badge, were only introduced as classes in September 2009 and September 2014 respectively. Panel B replaces the binary LEZ indicator with three indicators for each of the phases to zoom into the effect of each phase separately. The results suggest that the largest reduction in Euro 1 cars happens in Phase 2 and 3. The share of Euro 4 cars increases most strongly in Phase 3. These results are line with Wolff (2014), who also finds that LEZs create an incentive for drivers to substitute towards lower emitting vehicles.

3.5 The Effect of LEZs on Cardiovascular Health

This section presents the main findings of the paper. To analyze the outpatient data I estimate the following regression specification:

$$log(y_{at}) = \alpha + \beta LEZ_{at} + transition_{at} + T_t + D_a + \epsilon_{at}, \tag{3.3}$$

where the outcome is the number of patients with the given diagnosis in logarithms in area a in year t. LEZ_{at} indicates whether the area a has LEZ in January of year t. While most cities enact their LEZs at the beginning of the year, some introduce theirs later. Thus, I separate those periods by including the dummy *transition*

[17]Figure 3-12 in the appendix illustrates this development for each city separately.

that is one for all years when the zone was active for less than 12 months. T_t and D_a are year and area fixed effects. In this equation, area refers to either an LEZ, the peripheries of treated cities outside LEZs or an untreated city. Including area fixed effects accounts for time-invariant differences and ensures that identifying variation comes from within-area variation over time. It also takes into account general differences in size of the areas. Year fixed effects net out common time shocks and trends that affect all areas similarly.

As section 3.3 discusses, the main caveat of the outpatient data is missing observations before 2009, while the rollout of LEZs started already in 2008. This data restriction is particularly meaningful if the effect of LEZs on health is time varying. In a recent working paper, Goodman-Bacon (2018) shows that in difference-in-differences framework with timing variation, every unit acts as a control unit at some point. Hence, treatment effects that pick up over time in the post-treatment period will lead to a downward bias in the difference-in-differences coefficient. Therefore, I also present the results based on the sample that excludes previously treated cities from the estimation. Section 3.5.2 shows the event dynamics of the coefficients and implements the decomposition proposed in Goodman-Bacon (2018). Section 5.3 further juxtaposes the pollution and health effects in this restricted sample and provides more insights into the timing of the effects.

3.5.1 Main Results

Table 3.5 presents the main results for cardiovascular disease. Columns with odd numbers present estimates from the full sample. Columns with even numbers present the estimates from the restricted sample. Hereby, the treatment sample includes those cities that introduce LEZs after 2009.

Panel A presents the effect of LEZ on all cardiovascular disease. In column (1) the number of patients decreases by 1.2 percent, however the coefficient is imprecisely estimated. After restricting the sample, column (2) shows that the number of patients with cardiovascular disease of all ages decreases by 2.2 percent over the post-treatment period. This reduction translates into 0.9 fewer patients yearly per 10,000 people. Columns (3) and (4) present the effect for the elderly above 65. The point estimates suggest a tow to three percent reduction in the number of patients. This translates into approximately ten fewer patients yearly per 10,000 elderly.

In Panel B and C, I further slice the cardiovascular diagnoses into two subgroups: heart disease (I20-I49) and cerebrovascular disease (I60-I66). I focus on these two subgroups, as they are the most common cardiovascular diseases, accounting for a large share of costs, morbidity and mortality (Leal et al., 2006; Restrepo and Rieger, 2016). Together heart disease and cerebrovascular disease account for around 80 percent of all cardiovascular disease in the sample.

The estimates suggest that the decrease in heart disease is statistically significant only in the restricted sample. Here the effects are pronounced for those over the age of 65. The effects on cerebrovascular disease are much larger in magnitude and suggest a reduction of 7-13 percent both for the overall population and the elderly.

To compare the effect of LEZs across age groups, Figure 3-6 plots the coefficient from the restricted sample for age groups 15-29 and 30-64, juxtaposing these coefficients to the effect for population over 65. The prevalence of cardiovascular disease increases strongly with age. The fraction of people with a cardiovascular diagnosis is very low for those below the age of 30 (two percent at most). The prevalence increases in the age group 30-64, and and is rather high for those over the age of 65. The prevalence of heart disease and cerebrovascular disease is noticeably different. For example, 6.8 percent of 30-64 year olds have a heart disease diagnosis, but only 1.4 percent have a cerebrovascular disease. The pattern is similar for the elderly, albeit the respective numbers–42.4 percent and 11.1 percent–are much larger.

The figure shows that the number of all cardiovascular diagnoses decrease for all age groups, however, the largest reductions are indeed observed for the elderly patients. The pattern is different when separating the diagnosis group into heart disease and cerebrovascular disease. The estimates suggest that heart disease improves only for the elderly, while cerebrovascular disease improves for middle-aged adults as well.[18]

Considering prior work, the magnitude of the effects appears plausible. The German Environment Agency reports that multiple modeling studies initially suggested that LEZs can potentially reduce yearly average concentrations of PM_{10} by two percent in Phase 1 up to ten percent in Phase 3 compared to pollution concentrations in 2007 (Diegmann et al., 2007). I estimate a 1.1 $\mu g/m^3$, which is a four percent decrease in PM_{10} concentrations annually compared to 2007. This number, although closer to the lower bound, is within the interval of initial expectations.

Benchmarking the effect of LEZs on cardiovascular disease is not straightforward as, to the best of my knowledge, there are no quasi-experimental studies that link PM_{10} directly to cardiovascular disease. Hence, I rely on a prospective cohort study and a meta-analysis in 11 European cohorts by Cesaroni et al. (2012). The authors show a 10 $\mu g/m^3$ increase in annual mean PM_{10} is associated with 12 percent increase in cardiovascular events. The association persists also at low levels of exposure. A linear extrapolation implies that for a 1.1 $\mu g/m^3$ decrease in PM_{10} I should expect approximately a 1.3 percent drop in cardiovascular events. Estimates in Table 3.5 suggest a two percent reduction in cardiovascular events in general population and a three percent decrease for the elderly. My estimates are thus slightly larger, which might indicate that prior work underestimated the effect of PM_{10} on cardiovascular outcomes. Previous studies also suggest that the

[18] I cannot reject the null hypothesis that the two coefficients are equal to each other. The age-specific effects are not thoroughly discussed in the literature. Many epidemiological studies focus on the older individuals (most often over age 65) and exclude younger populations from the beginning. Studies that stratify the analysis by age groups find contradicting results. Some do in fact conclude that age has a strong effect on cerebrovascular health (Linn et al., 2000; Hong et al., 2002; Zhang et al., 2011; Maheswaran et al., 2012), while others find effects independent of age (Lee et al., 2018; Chau and Wang, 2020)

risk due to 1 $\mu g/m^3$ increase in PM_{10} is roughly comparable to the excess cardiovascular risk of smoking one cigarette a day, which is around four percent (Law et al., 1997).

A back-of-the-envelope cost-benefit analysis suggests that the benefits generated by the lower number of patients with cardiovascular disease surpass the costs that accrued due to induced vehicle upgrade by around two billion euros. Appendix C provides the details of the calculation.

3.5.2 Event Study and Goodman-Bacon Decomposition

The estimated coefficients in Table 3.5 report the effect of LEZ averaged over the entire study period. However, the LEZ-induced pollution reductions may have a lagged effect on health that may also change over time. Hence, I estimate an event-study model as follows:

$$log(y_{at}) = \alpha + \sum_{k=-3,k\neq-1}^{4} \beta_k LEZ_{ak} + T_t + D_a + \epsilon_{at} \quad (3.4)$$

Figure 3-5 presents the event study graphs for all cardiovascular diagnoses for all ages and the elderly over 65 separately. Each panel presents the coefficients and their 95% confidence intervals from the entire sample (red line) and after restricting the sample to cities treated after 2009 (blue line). The left panel plots the results for the entire population, and the right panel–for the elderly. Both graphs show that the trends in cardiovascular disease between treated and untreated cities display no clear trend before the implementation of LEZs. The pre-treatment coefficients are also jointly statistically insignificant. The respective p-values of the F-test for the hypothesis that all pre-treatment coefficients are jointly zero are 0.24 (entire sample) and 0.15 (introduction after 2009) in the left panel, and 0.14 for both samples in the right panel. Upon LEZs' enactment, the number of patients falls in treated cities at a faster rate than in untreated cities. Both graphs also show that the treatment effects are not time-constant: their absolute magnitude tends to become larger the longer LEZs are in place.

Time-varying treatment effects might bias the difference-in-differences estimate away from the true effect, as discussed in Goodman-Bacon (2018). The author shows that in difference-in-differences designs with timing variation, the difference-in-differences regression coefficient is a weighted average of all possible difference-in-differences coefficients of two-group two-period (2x2) comparisons, where the weights depend on the sample share and the treatment variance in each pair. It is possible to decompose and visualize each of these 2x2 estimates against their weight. The decomposition illustrates how average estimates vary across types of comparisons and which comparisons matter most.

To illustrate the proposed decomposition, Figure 3-7 plots the difference-in-differences estimates for cardiovascular health in the present setting. The graphs refer to all cardiovascular diagnoses from the entire sample for the overall population and the elderly. The vertical axis plots the 2x2 estimate for each pair and the horizontal axis plots the weight each of these pairs receive. The horizontal line shows the weighted average of all difference-in-differences estimates. The figure highlights the influential role of the pair "treatment versus never treated". In both panels, 68 percent of the variation comes from this comparison. This is not coincidental, as the variation share reflects the sample shares and the treatment variance, which are identical for both outcomes. The pure timing group comparisons get very small weights (2.5 percent for "earlier group treatment versus later group control", and 6.4 percent for "later group treatment vs earlier group control"). The figure also illustrates that the regression difference-in-differences coefficients might be smaller in magnitude due to time-varying effects. As the hollow circles show, the treatment versus already treated comparison mostly generates positive difference-in-differences coefficients with non-negligible weight (23.6 percent). It is possible to take out the bias from time-varying effects by subtracting the weighted average of all 2x2 difference-in-differences comparisons where the controls are the already treated units. This provides insight into why the difference-in-differences coefficients in Table 3.5 become larger in magnitude after restricting the treatment sample to cities that introduce LEZs after 2009. This empirical exercise supports the main findings and motivates the restricted specification in column (2) of Table 3.5 as the preferred specification.

3.5.3 The Timing of Pollution and Health Effects

This section examines the timing of changes in pollution and cardiovascular disease by unpacking the event dynamics and the separate effects of three LEZ phases.

The analysis of health data shows that the effects are strongest when the sample is restricted to treated cities that have introduced LEZs after 2009, as Figure 3-5 shows. Here I use the same restricted sample to demonstrate that in this subsample as well LEZs reduce air pollution. Furthermore, to gain insights into the timing of pollution and health effects, Figure 3-8 draws the event dynamics in this subsample for PM_{10}, NO_2 and cardiovascular disease.

There are two main takeaways from the event plots in Figure 3-8. First, LEZs reduce PM_{10} also in the subsample of post-2009 LEZ cities, whereas the reductions in NO_2 remain imprecisely estimated. Second, the reductions in the number of patients with cardiovascular diagnoses and in PM_{10} are congruent. The graph suggests that both PM_{10} and cardiovascular disease respond already in the implementation year of LEZs. This is not surprising. A large number of studies show that the impact of air pollution on health can be observed both in the short term and in the long term. Studies evaluating short-term variations in air pollution show that cardiovascular admissions and deaths respond already within a few hours up to two days (Pope, 2000; Wellenius et al., 2005; Fiordelisi et al., 2017).

The implementation of LEZs included several phases, as Section 3.2 discusses. These phases may affect the concentration of pollutants, and hence health outcomes differently. To separate the effects of the different phases, I replace the LEZ indicator with three indicators for each of the phases in equations (3.1) and (3.3), such that each indicator captures the effect of that phase individually. Figure 3-9 illustrates the coefficients and confidence intervals of these three phases. To juxtapose pollution and health effects, each graph combines the cardiovascular outcome under investigation with PM_{10} and NO_2. As in Table 3.5, I look at heart disease and cerebrovascular disease, as well as the elderly above 65 separately.

The results suggest that Phase 2 has the strongest impact on both pollutants. However, while the decreases of PM_{10} in Phase 2 and Phase 3 are comparable, the decrease in NO_2 is large in Phase 2, but bounces back in Phase 3. This result squares with two potential explanations. It is possible that Phase 2 achieves the largest environmental impact, as the number of yellow sticker cars on the roads is large. It is also possible that in this phase in most cities there were less cars on the roads, if the vehicle fleet has not managed to fully adapt yet. The pattern is different for health outcomes: the decrease in the number of patients becomes slightly larger in each phase. This development is suggestive evidence for a buildup in health effects.

3.5.4 Further Health Outcomes

A careful reading of epidemiological and economic literature suggests that respiratory disease, diabetes and central nervous system (CNS) disease are also frequently linked to air pollution. Economic literature has extensively investigated respiratory disease in particular. Most papers show a negative relationship between air pollution and respiratory health, especially for children (Lleras-Muney, 2010; Beatty and Shimshack, 2011; Moretti and Neidell, 2011; Janke, 2014; Schlenker and Walker, 2016; Simeonova et al., 2018; Alexander and Schwandt, 2019). Columns (1) and (2) in Table 3.16 in the appendix show that LEZs in Germany have no statistically significant effect on respiratory disease.[19] The results on respiratory health are somewhat at odds with the vast body of literature suggesting a strong

[19]The ICD-10 codes for respiratory disease are J00-J06, J20-J22, J30-J47 and J95-J99.

link between pollution and respiratory health.[20] To be able to completely rule out any impact on respiratory health, it would be helpful to analyze medical prescription data as well. Nevertheless, this discrepancy may largely be explained by institutional specificity of Germany. In 2006, Germany has implemented a disease management program for chronic respiratory illnesses, particularly COPD and asthma. Patients are enrolled in these programs by their health practitioner, who decides on a regular and fixed check-up interval (either quarterly or half-yearly). Health insurance companies provide monetary incentives to physicians to include patients in these programs (Mehring et al., 2013; Steppuhn et al., 2016). If reductions in pollution are not large enough to revert or prevent disease, it might be statistically challenging to detect changes in the number of patients and visits in such an institutional environment. Multiple studies show that air pollution affects cognitive function and central nervous system, especially dementia (Power et al., 2016; Babadjouni et al., 2017; Calderón-Garcidueñas and Villarreal-Ríos, 2017; Bishop et al., 2018; Peters et al., 2019) and diabetes (Andersen et al., 2012; Thiering and Heinrich, 2015; Eze et al., 2015). Table 3.16 shows the results for the diseases of the central nervous system, including dementia, in columns (3) and (4) and for diabetes in columns (5) and (6).[21] Across all specifications, the results suggest that the LEZ-induced reductions in air pollution do not have a large and statistically significant effect on these two outcomes. Disease management programs also exist for diabetes, hence, the lack of effect here as well might be specific to Germany's institutional setting.

3.5.5 Additional Results: Hospital Diagnosis Statistics

I supplement the evidence from outpatient data with inpatient data from the hospital admission records. The inpatient data offers the advantage of encompassing

[20] The respective result for children 0-6 is 0.010(0.017). Thus, there is no significant effect on children's respiratory health either.

[21] The respective ICD-10 codes are G00-G32, F00-09 and E10-E14.

a longer time period before the rollout of the zones. To construct the outcome variable, I count the number of episodes with cardiovascular diagnosis for each calendar year for the entire population and for those over the age of 65 separately. To evaluate the effect of LEZs on health using hospital data, I estimate a variant of equation (3.3), where the outcome is the number of hospital admissions with the given diagnosis per 10,000 population in logarithms in city c in year t. I further add time-varying controls at the city level, that is, GDP per capita, unemployment rate, average age of the population and the number of deceased. Table 3.6 presents the results. All estimates point to reductions in hospital admissions with cardiovascular disease. However, none of the estimates is precisely estimated at the 95% significance level. [22]

Two shortcomings of hospital data might explain this imprecision. First, the data contains no information on whether the person lives inside or outside an LEZ. Second, the inpatient data includes the cases that end up in hospitals, hence while it is ideal for studying severe cases, it is less suited for studying cases that can be managed by outpatient care.

3.6 Identification Threats and Robustness Checks

This section presents balancing regressions as a test of orthogonality and shows that the main results are robust to using a different control group, to sample restrictions, to alternative and placebo specifications, to different definitions of outcome and treatment variables.[23]

Balancing regressions

One potential threat to the identification strategy is that the implementation of LEZs might be correlated with simultaneous socioeconomic changes that affect pol-

[22]Figure 3-10 in the appendix presents the event study graphs for the hospital diagnosis statistics.

[23]Table 3.17 in the appendix presents the robustness checks for air pollution results.

lution and health outcomes. To test for such violations, I run balancing regressions (Pei et al., 2019; Alexander and Schwandt, 2019). In the spirit of difference-in-differences strategy, I regress the socioeconomic characteristics of treated areas on the LEZ indicator, along with city and year fixed effects. For each city, I collect data on industrial output per capita, unemployment rate, GDP per capita, output of health and other services and population density for the years 2004-2016 from the Federal Statistical Office. Panel A in Table 3.7 shows the corresponding results. Reassuringly, the coefficients on the LEZ indicator are insignificant in all regressions.

In Panel B I obtain two additional datasets from the Research Data Centre Ruhr at the Leibniz-Institute for Economic Research (FDZ Ruhr). The first dataset–RWI-GEO-GRID–contains a range of socioeconomic variables collected by the commercial micro- and geo-marketing provider *microm*, which uses more than one billion individual data points for aggregation. The data is available for the year 2005 and 2009-2017. From this dataset, I use variables on purchasing power per household,[24] population density, and the share of adults of the age 30-65 and over the age of 65. The second dataset–RWI-GEO-RED (RWI-GEO-RED)–is a unique dataset on German real estate prices, obtained from the largest internet platform on real estate in Germany. The data covers the years 2007-2017. From this dataset I use offering monthly rental prices for flats.

Both the socioeconomic variables and the rental prices are available at the postal code level. For consistency, I aggregate them at the area level, as defined in Section 3.5. Columns (6) through (10) show that there is no evidence that household income, age composition, population density, and rental prices change systematically differently within LEZs.

Definition of the control group

One may worry that using entire cities as a control group for LEZs may have

[24]Purchasing power is calculated to reflect household income. It includes labor income, capital wealth, rental and leasing income minus taxes and social security contributions.

implications for the estimates. LEZs are primarily placed in the densely populated central areas of major cities. Comparing these areas to the entire area of untreated cities might result in biased estimates if city centers systematically follow different time trends in health outcomes.

To address this concern, I generate hypothetical LEZs in untreated cities. Although municipal environmental agencies in each city follow an individual approach in placing their LEZs, one common predictor across all cities is the population density. The higher the density, the more likely it is that a postal code area is included in the LEZ. Hence, in untreated cities I draw the hypothetical LEZs by encircling the most densely populated postal code areas. For details on how the hypothetical LEZs are constructed, see Appendix B.

Table 3.8 presents the results from a regression akin to specification (3.3), however here the sample includes the existing and hypothetical LEZs only. The results support the main findings, suggesting that there is no evidence that city centers systematically follow different trajectories, despite differing in their socioeconomic characteristics at the baseline.

Specification issues

A potential concern is that the LEZs are a last resort policy for treated cities, which have enacted various anti-pollution policies previously. If those policies already beforehand put the treated cities on a different trajectory with respect to air pollution and health, the common trend assumption might be violated. However, the examination of the environmental plans of all cities included in the estimation sample shows that both treated and untreated cities have implemented various measures before the initial LEZs were enacted. These measures focused mainly on the expansion of public transport and cycling, retrofitting municipal vehicles, enacting parking policies and optimizing traffic flow. Since these policies were in place both in treated and untreated cities, they should affect all cities similarly. To further ensure that the results are not driven by differences in trends between treated and untreated cities, column (1) in Table 3.9 restricts the estimation sample to cities that have introduced an LEZ at some point during the observation period, meaning that the variation comes from the timing of the policy only. This robustness check also addresses the worry that treated and untreated cities are significantly different in their baseline characteristics, and hence might also systematically follow different trends. Reassuringly, the point estimates remain comparable to the main results.

Relatedly, the implementation of anti-pollution measures in untreated cities, if successful in reducing pollution, will lead to a downward bias in difference-in-differences estimates. Thus in column (2), I add a contemporaneous indicator variable for Clean Air Action Plans in untreated cities. The results show that the point estimates remain similar to the baseline results. Wolff (2014) also shows that there is little indication that other elements of Clean Air Action Plans have been effective.

Alternative outcomes

In the next robustness checks, I address the sensitivity of results with respect to the definition of the outcome. In column (3) of Table 3.9, the outcome is the number of cases instead of the number of patients, as in the main specification. This definition of the outcome should capture the intensive margin of the treatment effect as it also includes multiple visits by the same patient. The results suggest one percentage point larger drop in the number of patient visits versus the number of patients, suggesting that the reduction in the number of exacerbations plays a larger role.

Column (4) redefines the outcome in terms of prevalence rates, whereby prevalence rate is the fraction of patients with cardiovascular diagnosis in the total number of patients with any diagnosis in the respective age group. The results remain comparable in magnitude, however, the coefficient for the elderly sample becomes imprecisely estimated.

Placebo tests
The next block in Table 3.9 probes the sensitivity of results to three placebo tests. In column (5) I use the number of patients with injuries (ICD-10 W00-X59) as a placebo outcome. In column (6), I generate the timing of the LEZ enactments randomly for all treated cities. Subsequently, I drop all observations after the true treatment date and regress the outcome on the randomly generated LEZ indicator. In column (7), I drop all existing LEZs and use the number of patients with cardiovascular diagnosis in hypothetical LEZs in untreated cities as the outcome. Reassuringly, all three placebo tests return a small and insignificant coefficient.

Data aggregation at the city level
In the next block of robustness checks, I aggregate observations at the city level. This level of aggregation helps to examine whether the improvements in cardiovascular health are in the entirety of treated cities or only within the boundaries of LEZs. Panel A in Table 3.10 presents the results from a specification equivalent to the main regression, with the difference that outcomes are aggregated at the city level. The point estimates are very small and statistically insignificant both for the entire population and for the elderly. This emphasizes the importance of defining the treatment precisely.

To examine the issue further, panel B in Table 3.10 replaces the binary LEZ indicator with the size of the zone, calculated as a percentage of the LEZ in relation to the entire area of the city. This number is restricted between zero and one. The point estimate suggests that when the relative size of the zone increases by ten percentage points, the number of patients with cardiovascular disease in treated cities decreases by five percent for the entire population and by eight percent for those over the age of 65. Panel C further adjusts the outcome to the population size by calculating the number of patients per 10,000 inhabitants in the respective age group. The estimates suggest around seven percent reduction in cardiovascular diagnoses per 10,000 inhabitants for a ten percentage points increase in the relative size of the zone.

3.7 Conclusion

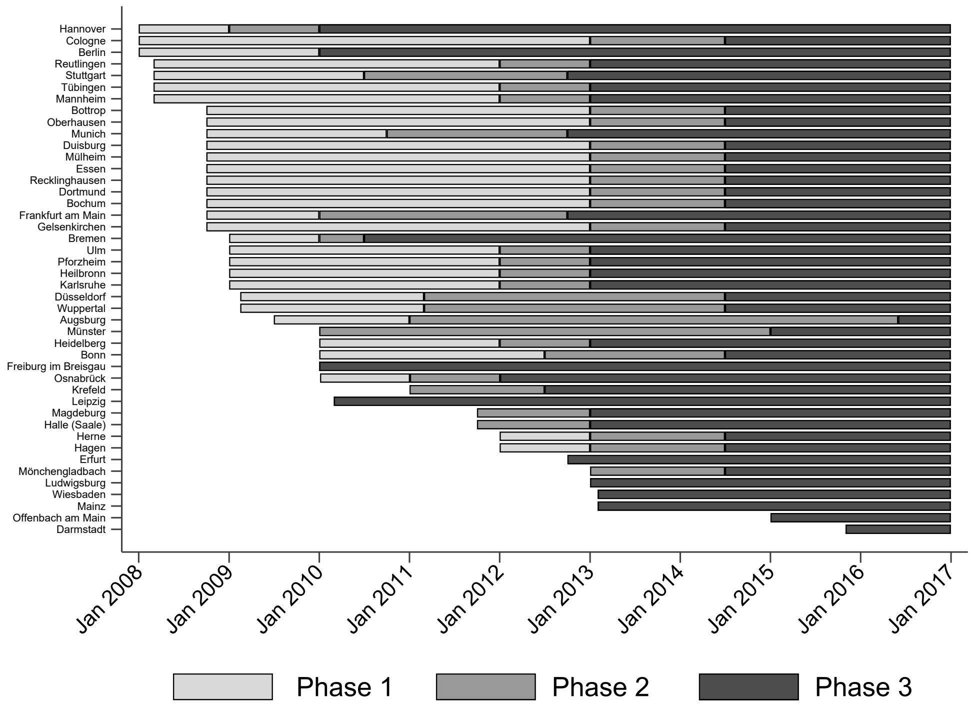
This paper evaluates the effect of LEZs on population health in Germany. Using the across space and over time variation in the implementation of LEZs, I first demonstrate that LEZs reduce monthly PM_{10} concentrations by 0.9 $\mu g/m^3$, which translates into a three percent decline. The findings for NO_2 suggest small and statistically insignificant reductions. Next, using novel outpatient health care data, I show that the zones improve cardiovascular health outcomes: they reduce the number of patients with cardiovascular disease by two to three percent. The effect is particularly strong for those over the age of 65. These results are robust to a range of robustness checks.

Despite the well-established understanding on the adverse effects of pollution, little empirical evidence exists on the instruments targeting air quality and their impact on health. The paper adds to this growing literature by providing evidence on the effect of LEZs. Furthermore, the findings in this paper show that adult population, especially those over the age of 65, being susceptible to cardiovascular health conditions, are an important group to study in the context of evaluating the effects of pollution on health. Finally, the paper adds to the small number of studies analyzing the effects of long-term marginal reductions in pollution in a setting with relatively low pollution levels and shows that, even at comparably low levels of air pollution, improvements in air quality generate health benefits.

The empirical findings in this paper have strong policy implications. They demonstrate that LEZs are a helpful tool to reduce air pollution in urban areas and to improve health outcomes commonly related to air pollution. A back-of-the-envelope cost-benefit analysis suggests health benefits of nearly 4.43 billion euros that have come at a cost of 2.3 billion euros for vehicle upgrading. However, the distributional effects of the policy are less clear. Since the policy mainly targets high-emitting cars, which tend to be old and cheap cars, the practical burden of LEZs most likely falls on families from low socioeconomic background and on small businesses. In the meantime, the car fleet has changed dramatically since the early introduction days of the zones. In 2019, the share of cars that received a green badge was above 90 percent. This weakens the impact of LEZs. Furthermore, the analysis in this paper indicates that the improvements in air quality, while meaningful, are not large enough to revert the existing health trajectories. If cities aim at reducing the air pollution further, stricter policy measures are necessary.

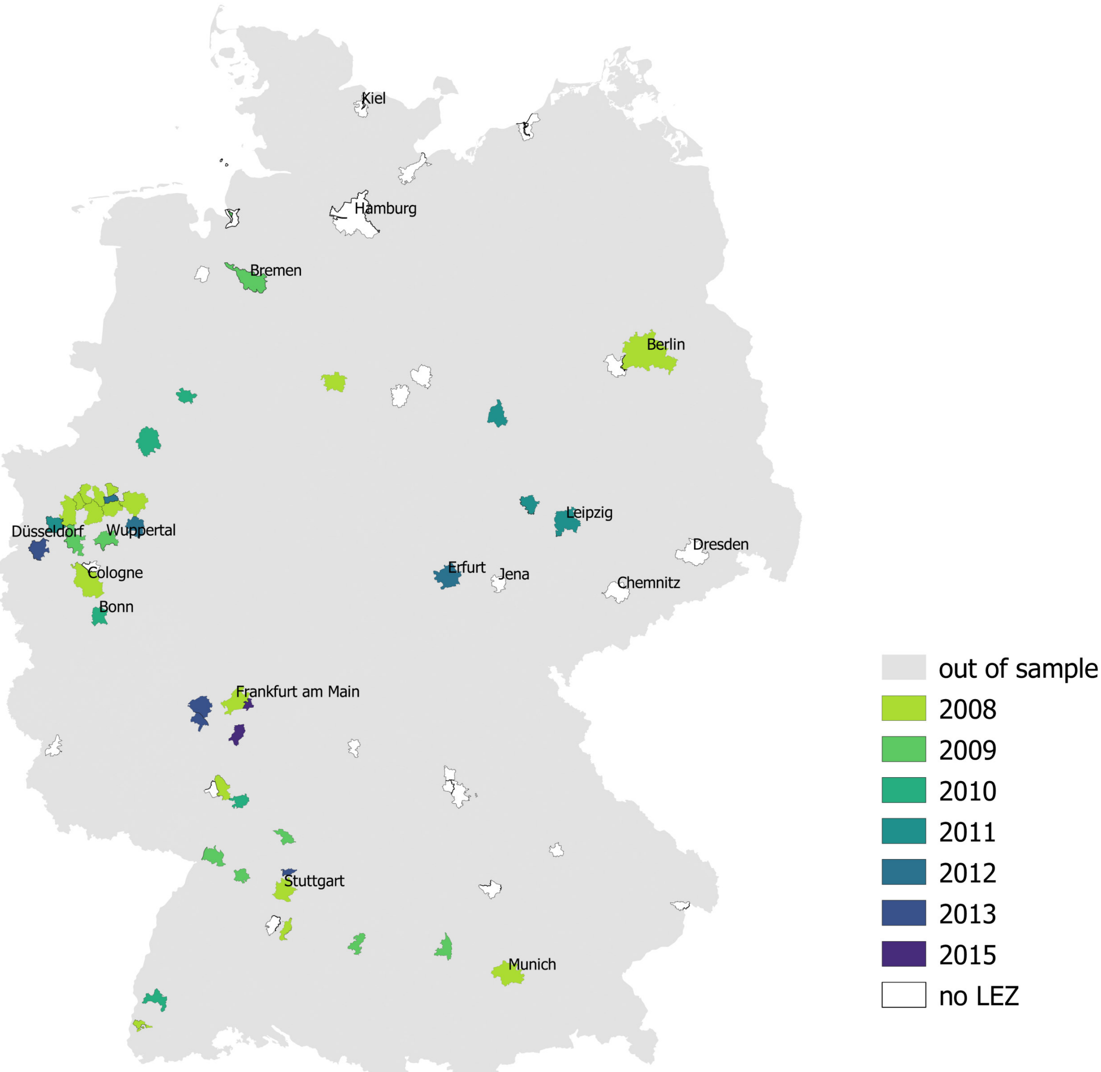
3.8 Figures and Tables

Figure 3-1: LEZ introduction and phases



Notes: The zones are introduced in three phases. Phase 1 restricts access to cars that receive no windscreen sticker. Phase 2 restricts access additionally to cars with red windscreen stickers, and Phase 3–additionally to yellow sticker cars. The color-coded stickers are given based on emission classes. Some of the cities, which introduced LEZs in 2011 or later, enacted zones directly as Phase 2 or Phase 3.

Figure 3-2: The variation in enactment of LEZ in major German cities

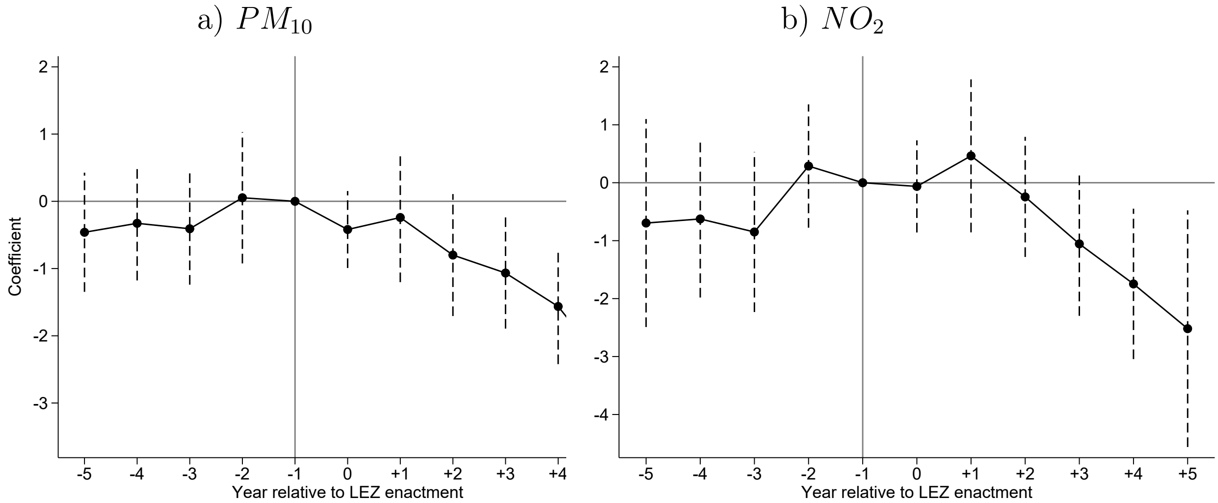


Notes: The map displays the untreated and treated cities, classified by the year of LEZ introduction. Some city names are displayed for ease of geographic orientation. Gray shaded areas are not part of estimation sample.

Figure 3-3: Berlin LEZ

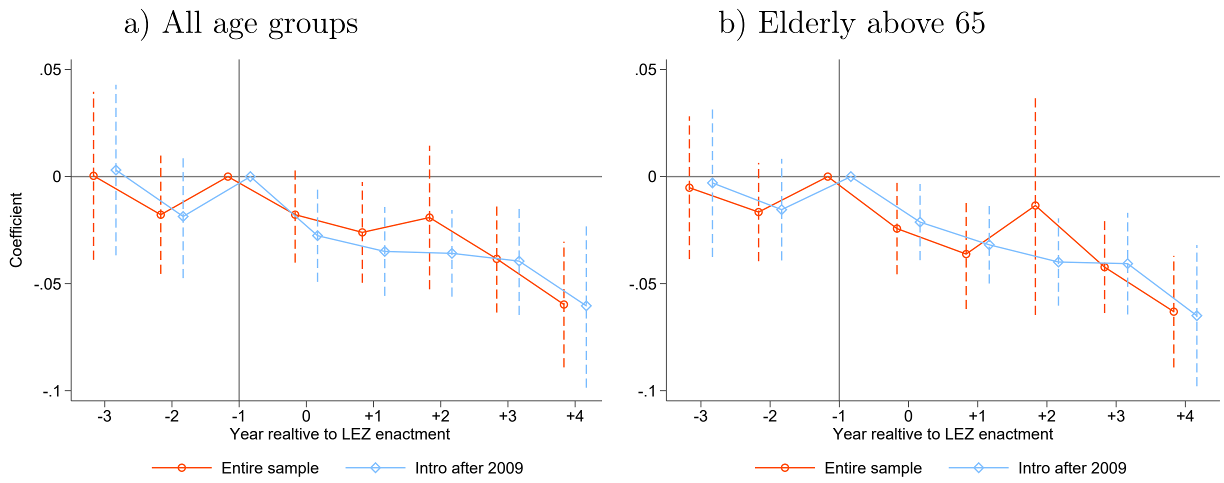
Notes: The map illustrates the Berlin LEZ (gray shaded) and the area outside (white shaded).

Figure 3-4: Event study of annual concentrations of PM_{10} and NO_2



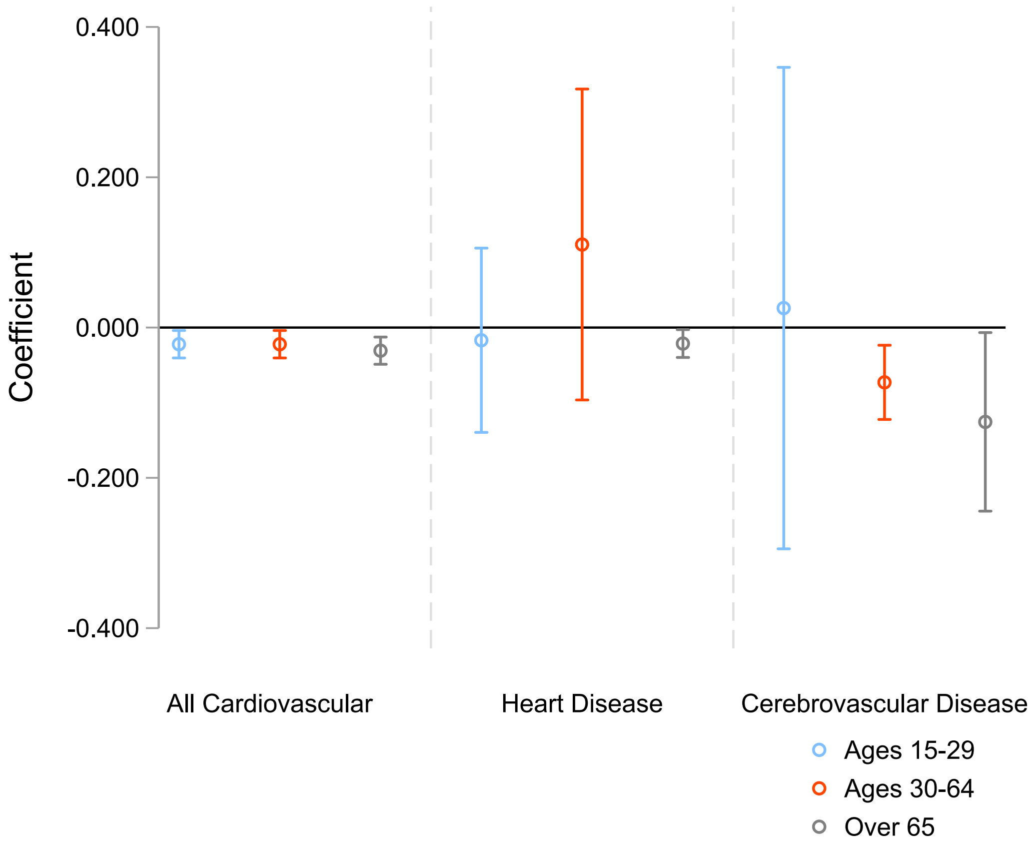
Notes: The outcome variable is the yearly average concentration of PM_{10} and NO_2. Coefficients and 95% confidence intervals from equation (3.2). The vertical line refers to t=-1 as the reference year. Standard errors are clustered at the city level. *Data:* German Environment Agency and German Weather Service.

Figure 3-5: Event study of cardiovascular disease



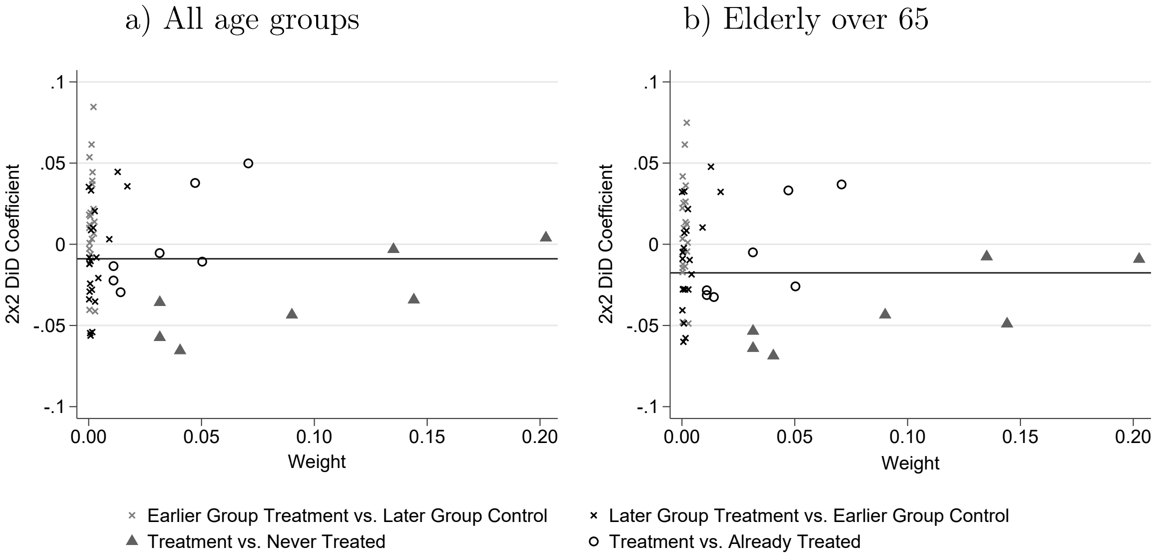
Notes: The outcome variable is the logarithm of the number of patients with cardiovascular disease. The red line reports the coefficients from a regression on the entire sample. The blue line restricts the sample to cities that have introduced an LEZ after 2009. Coefficients and 95% confidence intervals from equation (3.4). The vertical line refers to t=-1 as the reference year. Standard errors are clustered at the city level. *Data:* Zi.

Figure 3-6: The effect of LEZ on cardiovascular health, by sub-diagnoses and age groups.



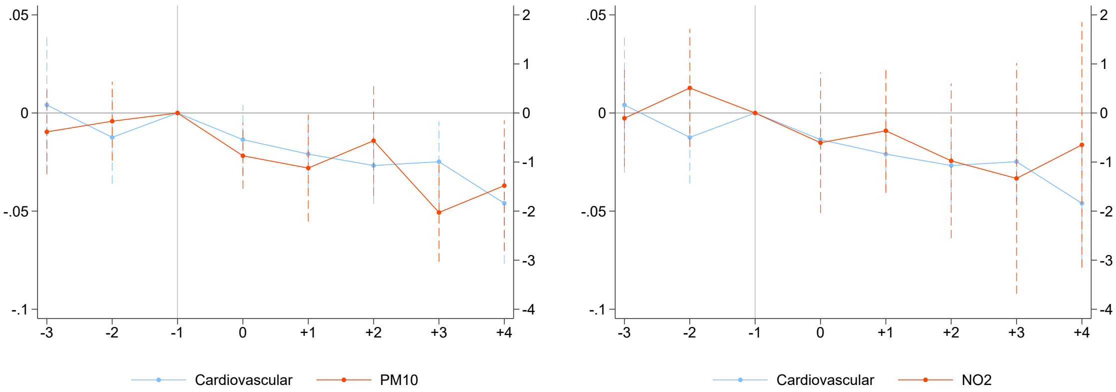
Notes: The model specification comes from equation (3.3). The figure presents results for three age groups from the subsample of cities that introduce LEZs after 2009. Coefficients and 95% confidence intervals. Robust standard errors are clustered at the city level. *Data:* Zi.

Figure 3-7: Difference-in-differences decomposition for cardiovascular disease



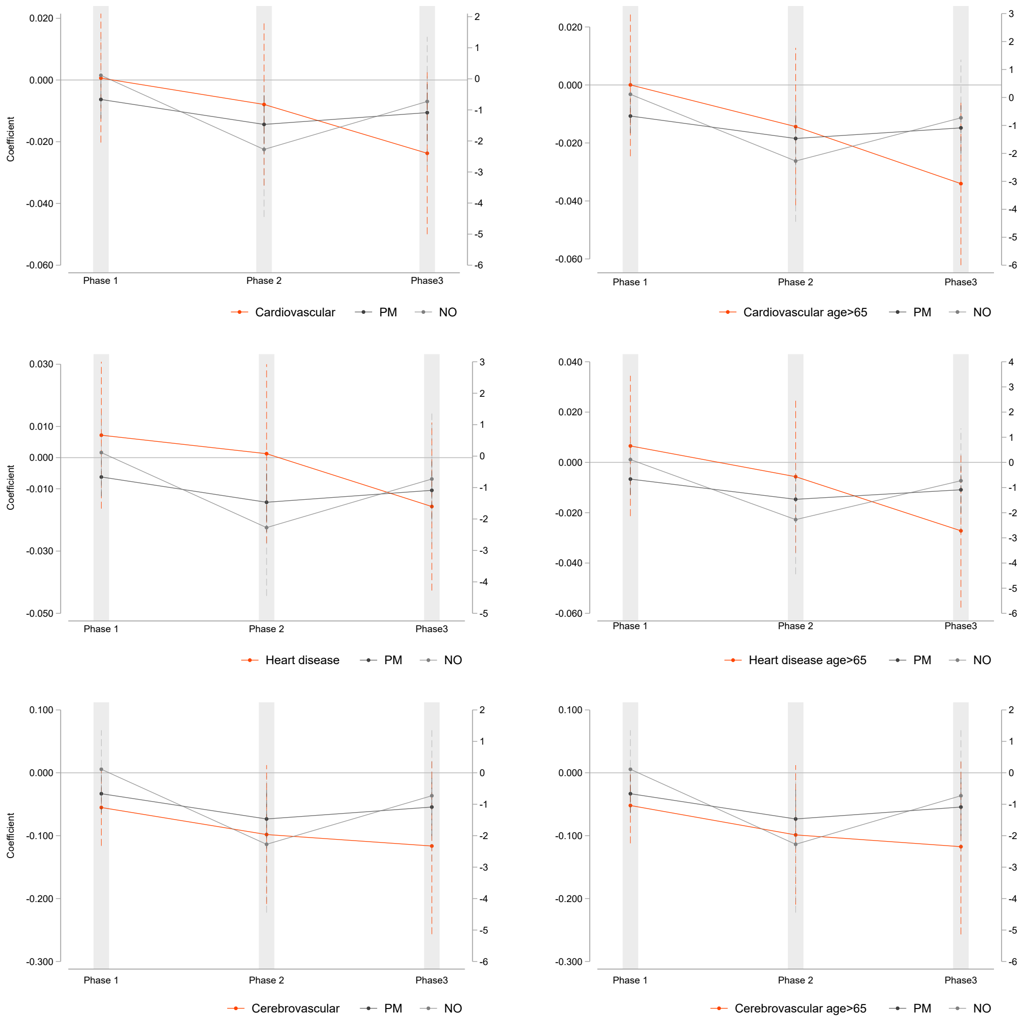
Notes: The figures plot each 2x2 difference-in-differences components from the Goodman-Bacon (2018) decomposition theorem. The horizontal line signifies the average DiD estimate, and equals to the sum of y-axis values weighted by x-axis values. *Data:* Zi

Figure 3-8: Event study of cardiovascular disease and annual concentrations of PM_{10} and NO_2



Notes: The figure presents coefficients from a subsample of treated cities that introduce LEZs after 2009. The vertical line refers to t=-1 as the reference year. Coefficients and 95% confidence intervals come from equations (3.2) and (3.4). Standard errors are clustered at the city level. *Data:* Zi, German Environmental Agency and German Weather Service.

Figure 3-9: LEZ phases and their impact on cardiovascular disease, PM_{10} and NO_2 separately



Notes: The figure presents coefficients from a subsample of treated cities that introduce LEZs after 2009. Coefficients and 95% confidence intervals come from equations akin to (3.1) and (3.3), where LEZ indicator is replaced with three indicators for each of the phases. Standard errors are clustered at the city level. *Data:* Zi, German Environmental Agency and German Weather Service.

Table 3.1: Limit Values for PM_{10} and NO_2 as defined by Council Directive 1999/30/EC

	Thresholds	Deadline
PM_{10}		
Yearly average limit	$40\mu g/m^3$	
Daily average limit	$50\mu g/m^3$	1 January 2005
Allowed number of transgression days per year	35	
NO_2		
Yearly average limit	$40\mu g/m^3$	
Daily average limit	$200\mu g/m^3$	1 January 2010
Allowed number of transgression days per year	18	

Notes: Source: Council Directive 1999/30/EC Annexes II, III.

Table 3.2: Means and Pre-trends of air quality and health outcomes

	PM_{10} (1)	NO_2 (2)	Cardiovascular (3)
	Average levels in 2007		
Untreated	23.4	34.08	9.51
Treated	27.02	42.23	9.31
p-value	0.00	0.00	0.87
	p-value on difference in pre-trends		
2005-2004	0.91	0.57	0.89
2006-2005	0.01	0.07	0.01
2007-2006	0.27	0.13	0.22

Notes: Variable names in headings. PM_{10} and NO_2 are measured in $\mu g/m^3$. Cardiovascular disease is measured as the number of admissions per 10,000 inhabitants. The lower panel presents the p-value on the difference in changes between treated and untreated cities. *Data*: German Environment Agency and Hospital Diagnosis Statistics.

Table 3.3: The effect of LEZ on monthly PM_{10} and NO_2 concentrations

	Monthly PM_{10}			Monthly NO_2		
	(1)	(2)	(3)	(4)	(5)	(6)
In LEZ	−0.872**	−0.622*	−0.869**	−0.818*	−0.748	−0.557
	(0.357)	(0.327)	(0.329)	(0.466)	(0.492)	(0.565)
Pre-LEZ mean	27.88	27.88	27.88	45.05	45.05	45.05
Change in percent	3.12	2.23	3.12	1.82	1.66	1.24
Observations	26,523	25,874	17,974	27,637	26,945	19,224
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	Yes	No	Yes	Yes
Station restrictions	No	No	Yes	No	No	Yes

Notes: The outcome is the monthly concentration of either PM_{10} or NO_2. Pre-LEZ mean refers to average concentrations of the respective pollutant at stations inside an LEZ before the enactment of the zone. Results from equation 1. Columns (1)-(2) and (4)-(5) include all stations. Columns (3) and (6) exclude stations in treated cities that are not inside an LEZ. Robust standard errors, clustered at the city level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* German Environment Agency and German Weather Service.

Table 3.4: The effect of LEZ on the share of cars by emission class

	Euro 1 (1)	Euro 2 (2)	Euro 3 (3)	Euro 4 (4)	Euro 4, 5 and 6 (5)
	Panel A: Single LEZ indicator				
LEZ	–0.354**	–0.329	0.119	1.891*	0.564
	(0.152)	(0.340)	(0.181)	(1.119)	(0.376)
	Panel B: Indicator for each phase				
Phase 1	–0.132	–0.154	–0.011	1.172	0.296
	(0.116)	(0.274)	(0.166)	(0.874)	(0.269)
Phase 2	–0.408**	–0.395	0.208	1.921	0.595
	(0.166)	(0.361)	(0.190)	(1.342)	(0.390)
Phase 3	–0.540**	–0.480	0.221	2.560*	0.799
	(0.216)	(0.432)	(0.250)	(1.368)	(0.534)
Observations	690	690	690	690	690
Year FE	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes

Notes: The outcome is the share of cars in each emission class in the entire registration of cars annually in a given city. Panel A controls for a single binary indicator for LEZ. Panel B includes binary indicators for each of the phases. Robust standard errors, clustered at the city level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Federal Motor Vehicle Authority.

Table 3.5: The effect of LEZ on cardiovascular disease

	All age groups		Elderly above 65	
	(1)	(2)	(3)	(4)
	Panel A: All cardiovascular disease			
LEZ	–0.012	–0.022**	–0.021**	–0.031***
	(0.010)	(0.009)	(0.009)	(0.009)
Pre-LEZ mean	9.81	9.81	9.48	9.48
Observations	954	585	954	585
	Panel B: Heart disease			
LEZ	–0.006	–0.012	–0.016	–0.021**
	(0.010)	(0.009)	(0.011)	(0.009)
Pre-LEZ mean	9.25	9.25	8.88	8.88
Observations	954	585	954	585
	Panel C: Cerebrovascular disease			
LEZ	–0.072**	–0.126**	–0.071*	–0.126**
	(0.036)	(0.059)	(0.036)	(0.059)
Pre-LEZ mean	7.86	7.86	7.56	7.56
Observations	954	585	954	585
Area FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Intro after 2009	No	Yes	No	Yes

Notes: The outcome is number of patients with the given diagnosis in logarithms. The restriction "Intro after 2009" refers to keeping only those treated cities that introduced an LEZ after 2009. Pre-LEZ mean refers to the logarithm of average number of patients in the entire pre-treatment period across cities. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Zi.

Table 3.6: The effect of LEZ on cardiovascular disease. Hospital diagnosis statistics

	Entire population (1)	Elderly above 65 (2)
	Panel A: All cardiovascular disease	
LEZ	−0.037	−0.030
	(0.031)	(0.034)
Observations	726	726
	Panel B: Heart disease	
LEZ	−0.048	−0.046
	(0.035)	(0.036)
Observations	726	726
	Panel C: Cerebrovascular disease	
LEZ	−0.050	−0.042
	(0.052)	(0.058)
Observations	726	726
City FE	Yes	Yes
Year FE	Yes	Yes
Additional controls	Yes	Yes

Notes: The outcome is the number of patients with the given diagnosis per 10,000 population in logarithms. The additional controls are GDP per capita, unemployment rate, average age of the population and the number of deceased. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*Data:* Hospital Diagnosis Statistic.

Table 3.7: Balancing regressions

	Panel A: Characteristics at city level				
	Industrial output per capita (1)	Unemployment rate (2)	GDP per capita (3)	Health and services per capita (4)	Population density (5)
LEZ	0.215	0.001	–447.2	–0.027	8.184
	(0.243)	(0.029)	(311.6)	(0.108)	(13.62)
Observations	897	897	897	897	897
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
	Panel B: Characteristics at LEZ level				
	Purchasing power (6)	Share of age>65 (7)	Share of 30<age<65 (8)	Population density (9)	Rental prices (10)
LEZ	–110.627	–0.177	–0.091	–64.860	0.200
	(414.179)	(0.131)	(0.101)	(62.677)	(0.382)
Observations	1,290	1,290	1,290	1,290	719
Area FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variables in column headings. In Panel A, LEZ indicates whether a city has a low emission zone at a given time. In Panel B, LEZ indicates whether a given postal code is inside a low emission zone. Robust standard errors, clustered at the city level, in parentheses. *Data:* Federal Statistical Office of Germany (Panel A), RWI-GEO-GRID and RWI-GEO-RED (Panel B).

Table 3.8: The effect of LEZ on cardiovascular disease.
Existing and hypothetical LEZs

	Entire population		Elderly above 65	
	(1)	(2)	(3)	(4)
	Panel A: All cardiovascular disease			
LEZ	−0.013	−0.015	−0.026*	−0.028*
	(0.017)	(0.018)	(0.015)	(0.016)
Observations	594	396	594	396
	Panel B: Heart disease			
LEZ	−0.003	−0.008	−0.012	−0.016
	(0.023)	(0.024)	(0.022)	(0.022)
Observations	594	396	594	396
	Panel C: Cerebrovascular disease			
LEZ	−0.048	−0.080**	−0.056*	−0.090***
	(0.029)	(0.030)	(0.029)	(0.029)
Observations	954	585	954	585
Area FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Intro after 2009	No	Yes	No	Yes

Notes: The outcome is number of patients with the given diagnosis in logarithms. Area FE are fixed effect referring either to existing or to hypothetical LEZs. The restriction "Intro after 2009" refers to keeping only those treated cities that introduced an LEZ after 2009. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Zi.

Table 3.9: Robustness Checks

	Specification issues		Alternative outcomes		Placebo tests		
	LEZ cities only	Clean Air Action Plans	Patient-cases	Prevalence rate	Pl. outcome: Injuries	Pl. timing	Hypothetical LEZs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Entire population	–0.024*	–0.036***	–0.033***	–0.033**	0.372	0.000	–0.005
	(0.012)	(0.008)	(0.009)	(0.012)	(0.393)	(0.012)	(0.011)
Elderly above 65	–0.030**	–0.043***	–0.041***	–0.015	0.018	–0.004	–0.006
	(0.012)	(0.009)	(0.009)	(0.012)	(0.124)	(0.011)	(0.013)
Observations	360	585	585	585	585	333	603
Area FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Intro after 2009	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Column (1) excludes all cities with no LEZ by 2017. Column (2) additionally controls for Clean Air Action Plans with an indicator variable. Column (3) and (4) use alternative outcomes. In column (3) the outcome is the overall number of patient-cases in logs. In column (4) the outcome is the number of cardiovascular patients per 10,000 patients with any diagnosis, in logs. Column (5) uses the number of patients with injuries as a the placebo outcome. Column (6) replaces the actual LEZ indicator with a randomly generated fake indicator. Column (7) uses hypothetical LEZs instead of existing LEZs. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Zi.

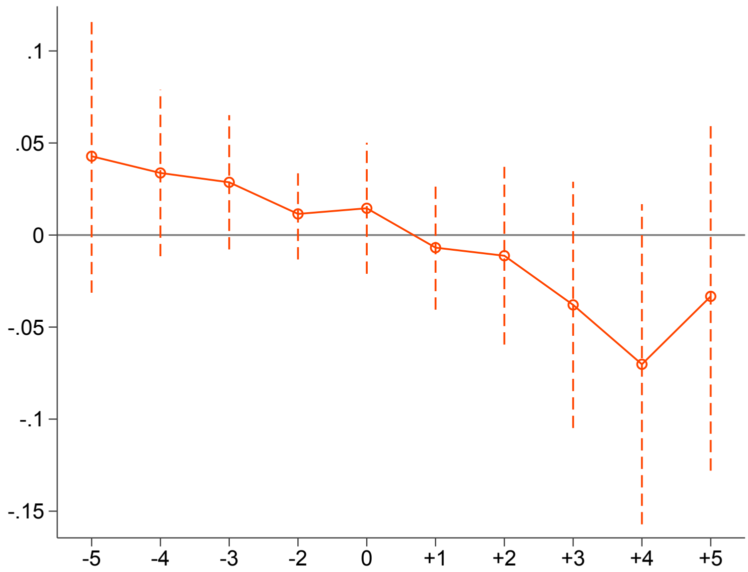
Table 3.10: Robustness Checks: Aggregation at the city level

	Entire Population	Elderly above 65
	(1)	(2)
	Panel A: LEZ indicator for the city	
LEZ	−0.009	−0.013
	(0.014)	(0.015)
Observations	396	396
	Panel B: Proportional size of the LEZs	
LEZ size	−0.049***	−0.054***
	(0.015)	(0.017)
Observations	396	396
	Panel C: Number of patients per 10000 inhabitants	
LEZ size	−0.077***	−0.050**
	(0.014)	(0.022)
Observations	396	396
City FE	Yes	Yes
Year fixed effects	Yes	Yes
Intro after 2009	Yes	Yes

Notes: In Panel A, LEZ indicator takes on the value one if the city has an LEZ, and zero otherwise. Panel B replaces the binary LEZ indicator with the size of the zone, calculated as a percentage of the LEZ in relation to the entire area of the city. Panel C further adjusts the outcome to the population size by calculating the number of patients per 10,000 inhabitants in the respective age group. Robust standard errors, clustered at the city level, in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. *Data:* Zi.

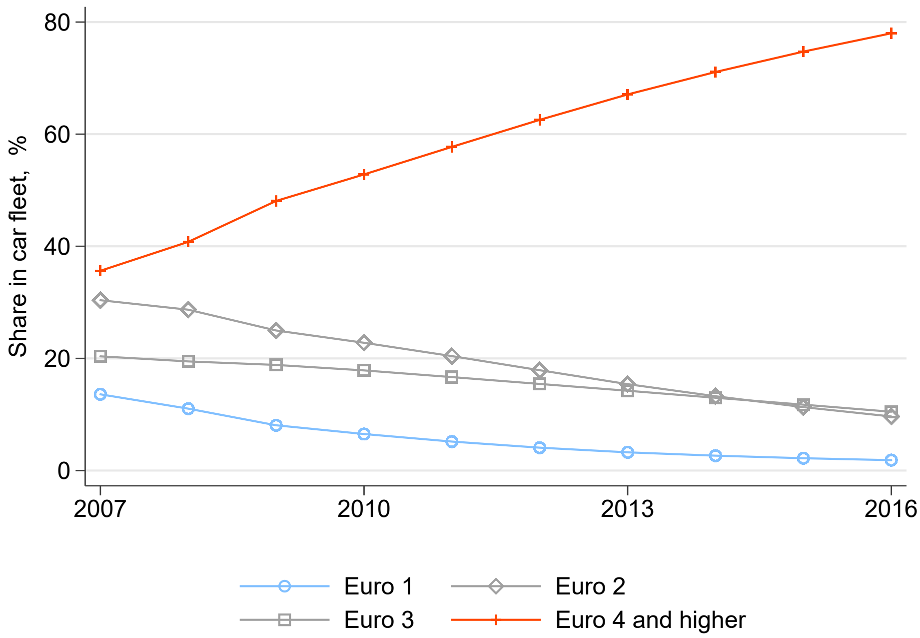
3.9 Appendix A: Additional Figures and Tables

Figure 3-10: Event study of cardiovascular diseases. Hospital data.



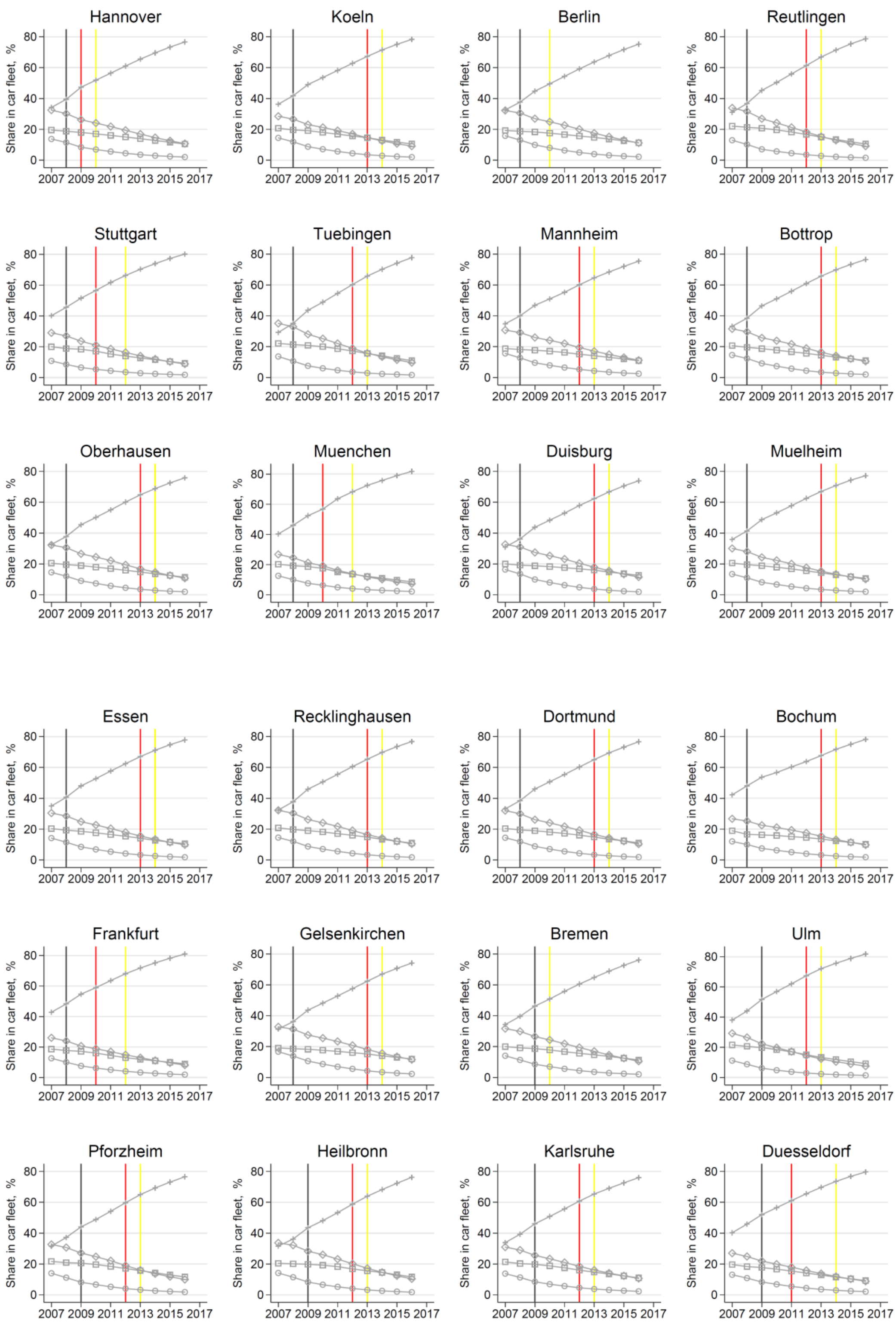
Notes: The outcome variable is the logarithm of the number of patients with cardiovascular disease per 10,000 population. Coefficients and 95% confidence intervals. *Data*: Hospital Diagnosis Statistics.

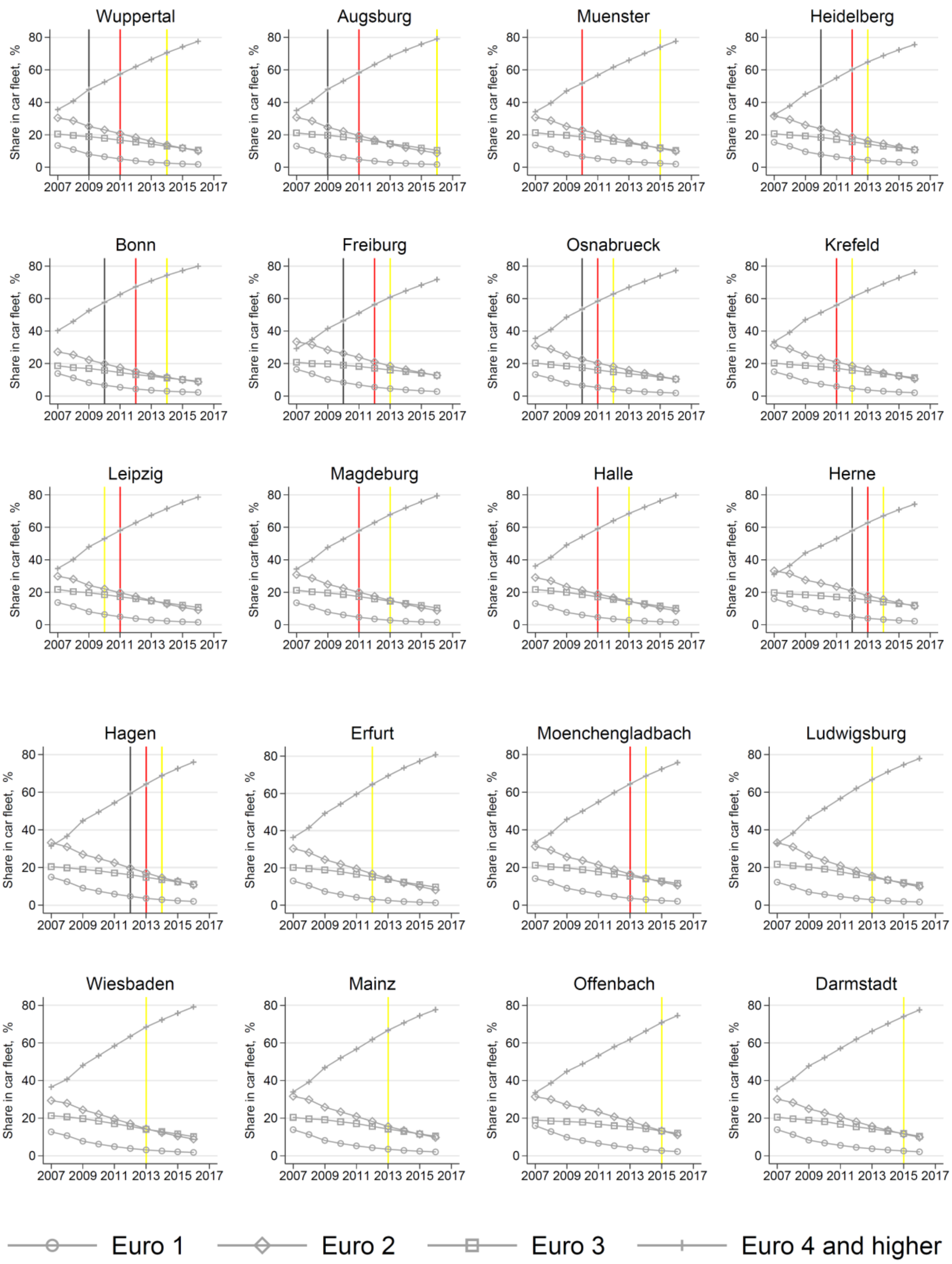
Figure 3-11: Development of passenger vehicle registration by emission class



Notes: Share of each Euro-class is calculated as its percentage in the total registration of cars annually in a city. *Data:* Federal Motor Vehicle Authority.

Figure 3-12: Development of passenger vehicle registration by emission class in treated cities





Notes: The figures plot the share of cars in each emission class for each treated city separately. The gray horizontal line signifies Phase 1, the red line signifies Phase 2 and the yellow line–Phase 3. *Data*: Federal Motor Transport Authority.

Table 3.11: Emission classes, color codes and phase restrictions

Emission class	**Colour code**	**Banned in**	**Description**
Euro 4	Green	None	Petrol: CO: 1.00g/km HC: 0.10g/km NOx: 0.08g/km Diesel: CO: 0.50g/km HC + NOx: 0.3g/km PM: 0.025g/km
Euro 3	Yellow	Phase 3	Petrol: CO: 2.30g/km HC: 0.20g/km NOx: 0.15g/km Diesel: CO: 0.64g/km HC: 0.56g/km NOx: 0.50g/km PM: 0.05g/km
Euro 2	Red	Phase 2	Petrol: CO: 2.20g/km HC + NOx: 0.50g/km Diesel: CO: 1.00g/km HC + NOx: 0.70g/km PM: 0.08g/km
Euro 1	None	Phase 1	Petrol CO: 2.72g/km HC + NOx: 0.97g/km Diesel: CO: 2.72g/km HC + NOx: 0.97g/km PM: 0.14g/km

Notes: The table displays European emission standards for exhaust emissions of new vehicles.

Table 3.12: Typical cars in Euro-class emission category classification

Euro 1	Euro 2	Euro 3	Euro 4 and higher
Volkswagen Caddy	Volkswagen Passat Variant	Nissan Primera	Volkswagen Tiguan
Mercedes-Benz E250	Mercedes-Benz C220	Mercedes-Benz Vaneo	Skoda Kodiaq
Audi 80 TDI	Citroen Xantia	Volvo V70	Mitsubishi Pajero
Volkswagen T4 California	Mercedes-Benz ML 270	BMW 520	Ford Fiesta
Volvo 940	BMW 740	Fiat Scudo	Ford Transit
		BMW 330	Volvo XC 60
		Opel Vivaro	Tesla
			Civic hybrid
			Prius

Notes: The table displays the typical cars in Euro classifications.

Table 3.13: Introduction date and the areal coverage of LEZs

City	Introduction	Coverage in %	Attainment Status
Augsburg	July 2009	3.95	non-attainment
Berlin	January 2008	9.88	non-attainment
Bochum[1]	October 2008	39.85	
Bonn	January 2010	6.38	attainment
Bottrop[1]	October 2008	24.85	non-attainment
Bremen	January 2009	2.18	
Cologne	January 2008	7.42	attainment
Darmstadt	November 2015	87.64	non-attainment
Dortmund[1]	October 2008	6.77	non-attainment
Duisburg[1]	October 2008	18.47	non-attainment
Dusseldorf	February 2009	19.78	
Erfurt	October 2012	5.86	non-attainment
Essen[1]	October 2008	66.56	non-attainment
Frankfurt	October 2008	44.3	non-attainment
Freiburg	January 2010	16.18	non-attainment
Gelsenkirchen[1]	October 2008	19.06	non-attainment
Hagen	January 2012	5.39	non-attainment
Halle	October 2011	5.1	non-attainment
Hannover	January 2008	21.05	non-attainment
Heidelberg	January 2010	9.29	attainment
Heilbronn	January 2009	38.34	non-attainment
Herne[1]	January 2012	100	
Karlsruhe	January 2009	6.52	non-attainment
Krefeld	January 2011	23.2	non-attainment
Leipzig	March 2011	61.35	non-attainment
Ludwigsburg	January 2013	100	non-attainment
Magdeburg	October 2011	3.33	non-attainment
Mainz	February 2013	34.95	non-attainment
Mannheim	March 2008	4.67	non-attainment
Mönchengladbach	January 2013	12.38	attainment
Mülheim[1]	October 2008		attainment
Munich	October 2008	14.16	
Münster	January 2010	0.47	attainment
Oberhausen[1]	October 2008	30.87	
Offenbach	January 2015	85.77	
Osnabrück	January 2010	14.11	attainment
Pforzheim	January 2009	1.99	non-attainment
Recklinghausen[1]	October 2008	30.08	
Reutlingen	March 2008	100	non-attainment
Stuttgart	March 2008	98.44	non-attainment
Ulm	January 2009	23.07	non-attainment
Wiesbaden	February 2013	31.12	attainment
Wuppertal	February 2009	14.61	non-attainment

Notes: Dates of introduction of the zones come from the *Umweltbundesamt*. The coverage refers to the relative size of LEZ in relation to the area of the city. The size of LEZ has been calculated based on shapefiles from *OpenStreetMap*.

[1] The cities in Ruhr area united into a common LEZ in January 2012.

Table 3.14: Differences between LEZ and non-LEZ areas in treated cities in 2005

	Purchasing power	Population density	Car density	Unemployment rate	Share of 30-65	Share of 65+
Mean treated	37,359	4,865	0.72	10.28	49.87	17.73
Mean untreated	39,671	1,899	0.94	8.40	49.44	19.45
Difference	-2,311	2,966	−0.217	1.888	0.429	−1.723
P-value	0.026	0.000	0.000	0.018	0.188	0.001

Notes: N treated=43 and N control=38. Purchasing power and car density are calculated per household. Populations density refers to number of people per km^2. *Data*: RWI-GEO-GRID.

Table 3.15: The effect of LEZ on monthly $PM_{2.5}$ concentrations

	(1)	(2)	(3)
LEZ	−0.435**	−0.332	−0.472*
	(0.205)	(0.210)	(0.253)
Observations	7,114	6,922	6,594
Station fixed effects	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes
Weather controls	No	Yes	Yes
Station restrictions	No	Yes	Yes

Notes: The outcome is the monthly concentration of $PM_{2.5}$. Column (1) includes all stations and does not control for weather covariates. Column (2) includes all stations and controls for weather covariates. Column (3) restricts the sample to stations that functioned at least six months before and after the introduction of the respective LEZ zone. Robust standard errors, clustered at the city level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* German Environment Agency and German Weather Service.

Table 3.16: The effect of LEZ on further outcomes

	Respiratory disease		**CNS disease**		**Diabetes**	
	All ages	Over 65	All ages	Over 65	All ages	Over 65
	(1)	(2)	(3)	(4)	(5)	(6)
LEZ	0.004	−0.065	−0.005	0.003	−0.004	−0.011
	(0.006)	(0.060)	(0.014)	(0.041)	(0.007)	(0.010)
Observations	585	585	585	585	585	585
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Intro after 2009	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The outcome is number of patients with the given diagnosis in logarithms. Area FE refers to fixed effect for each district over which the numbers are aggregated. The restriction "Intro after 2009" refers to keeping only those treated cities that introduced an LEZ after 2009. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Zi.

Table 3.17: The effect of LEZ on monthly PM_{10} and NO_2 concentrations. Robustness checks

	Monthly PM_{10}		Monthly NO_2	
	(1)	(2)	(3)	(4)
Panel A: Only LEZ Cities				
LEZ	−0.635**	−0.656*	−0.581	−1.014***
	(0.273)	(0.360)	(0.371)	(0.348)
Observations	13014	13014	13012	13012
Panel B: Restricted Observation Period				
B.1: Excluding 2004 and 2005				
LEZ	−0.796***	−0.749**	−0.562*	−1.053**
	(0.260)	(0.301)	(0.304)	(0.437)
Observations	16477	16477	17293	17293
B.2:Excluding 2013 and 2014				
LEZ	−0.693**	−0.859**	−0.389	−0.795**
	(0.322)	(0.328)	(0.368)	(0.354)
Observations	16357	18099	17039	17039
B.3: Excluding 2004-2005 and 2013-2014				
LEZ	−0.705**	−0.823**	−0.778**	−1.259***
	(0.308)	(0.340)	(0.296)	(0.345)
Observations	13025	13025	13528	13528
Panel C: Placebo Timing				
LEZ	0.363	0.132	0.327	0.406
	(0.372)	(0.381)	(0.601)	(0.696)
Observations	13562	13562	14542	14542
Station fixed effects	Yes	Yes	Yes	Yes
Year-month fixed effects	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Station restrictions	No	Yes	No	Yes

Notes: The outcome is the monthly concentration of PM_{10} and NO_2. Panel A restricts the sample to treated cities only. Panel B restricts the observation period. Panel C regressed the outcome on a randomly generated LEZ introduction date. Robust standard errors, clustered at the city level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data*: German Environment Agency and German Weather Service.

3.10 Appendix B: Construction of Hypothetical LEZs

Municipal environmental agencies in each city follow an individual approach in placing their LEZs. This generates a vast variation in the characteristics and proportional size of LEZs across treated cities. Nevertheless, across treated cities population density in a postal code is a common predictor whether that postal code is inside an LEZ. The higher the density, the more likely it is that a postal code area is included in the LEZ. Columns (1) and (2) in Table 3.18 show the strong and significant correlation between the probability that a postal code area is part of an LEZ and population density. A descriptive analysis also shows that treated postcode areas are at least in the top 50th percentile of population density distribution. Furthermore, all LEZs, at least partially, cover the center of the city.

To generate the hypothetical LEZs, in analogy to treated cities, I define the postal code areas that fall in the top 50th percentile according to population density in untreated cities. I define the density cutoffs for each city separately, instead of defining an absolute cutoff across the entire sample. Since there is a large between-city variation, an absolute density cutoff would lead to no potential postal codes for hypothetical LEZ construction in certain cities. On a choropleth map with postal codes shaded in proportion to population density, I manually define the hypothetical LEZs for untreated cities, encircling postal codes in the top 50th percentile of density distribution that fall in the city center. To ensure a contiguous region in the center of the city, I exclude the postal code areas that are not located in the city center, even when these fall in the top percentiles of density distribution.[25] Figure 3-13 also illustrates this exercise visually on the example of the map of Berlin (treated) and Hamburg (untreated). Columns (3) and (4) in Table 3.18 show the correlation between the probability that a postal code is included in the hypothetical LEZ and population density. As before, the strong and significant correlation remains.

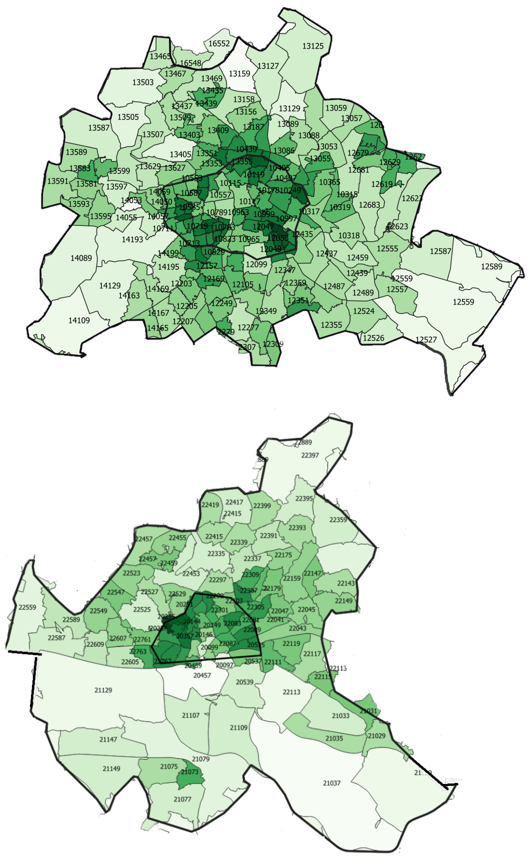
[25] An example of such a postal code area is postal code 21073 in Hamburg in Figure 3-13.

Table 3.18: LEZs and population density at the postal code level

	Existing LEZ		Hypothetical LEZ	
	(1)	(2)	(3)	(4)
Density	0.039***	0.055***	0.074***	0.099***
	(0.005)	(0.004)	(0.011)	(0.016)
Observations	923	923	345	345
R-squared	0.109	0.436	0.262	0.481
City FE	No	Yes	No	Yes

Notes: The outcome is binary and equals one if the postal code is inside an existing or hypothetical LEZ. Robust standard errors, clustered at the city level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. *Data:* Census 2011.

Figure 3-13: Population density by postal codes in Berlin and Hamburg



Notes: The figures show the population density per sq/km. The darker the shade, the more densely is the postal code populated. On the left is Berlin, with its existing LEZ contoured. On the right is Hamburg, with its hypothetical LEZ contoured.

3.11 Appendix C: Cost-Benefit Analysis

An official cost-benefit assessment, unfortunately, is not available. This section presents a cost-benefit assessment under multiple subjective assumptions. This analysis is intended to provide a rough direction rather than a precise assessment.

Calculating the cost of LEZs requires multiple arbitrary assumptions. Although various factors are part of the total cost, clearly, the largest costs are due to vehicle upgrade. Here the difficulty is that it is unclear which part of this upgrade would have happened in the absence of LEZs and which is directly the result of LEZs.

I calculate the costs of car fleet update (direct costs) following the strategy in Wolff (2014). Following the author, I assume that the average cost of upgrading a car is 1650 euros. First, I calculate the yearly percentage change in green sticker cars in treated cities. I then subtract the lowest average pace of fleet update in control cities (2.5 percent) to approximate the change in car fleet that is independent of the LEZs. I multiply this final percentage with the total number of cars and multiply this number by 1650. Finally, I sum up the costs over all treated cities in their respective post-treatment period. This results in total costs of 2.3 billion euros. It is worth noting that there are other cost factors that this number does not take into account, such as time lost due to taking longer routes or using other means of transportation (public transport, car sharing, cycling) and the potential negative effect of LEZs on (small) businesses inside the zone.

Calculating benefits generated by the LEZs is complicated as well since there is no official, publicly available cost estimate by patient age groups and diagnoses. Moreover, proper benefit calculations would require factoring in not only direct, but also indirect benefits, such as the utility to residents from low pollution. Keeping all shortcomings in mind, below I follow a simple approach in calculating the benefits generated by fewer number of cardiovascular diagnoses.

According to the cost of illness report of the German Statistical Office (Statistisches Bundesamt, 2015), per capita cost of illness of cardiovascular disease is 570 euros. Meanwhile, the per capita cost of cardiovascular disease for those over the age of 65 is 2600 euros. General cost breakdown suggests that around 50 percent of all healthcare costs are generated by the outpatient care. This implies that ambulatory healthcare per capita cost is 235 euros for overall population and 1300 euros for those over the age of 65. The Robert Koch Institute reports that 8.5 percent of overall population and 23 percent of those older than 65 are diagnosed with a cardiovascular disease. I use the latter statistic to calculate the costs per patient for general population and for the elderly. The calculation leads to 2764.7 euros savings for overall population (100/8.5*235) and 5652.2 euros for the elderly (100/23*1300):

My estimates in Section 3.5 suggest that LEZs reduced the number of patients with a cardiovascular diagnosis by 3.3 percent. Using these numbers, I calculate that the outpatient health care system saved 4,43 billion euros.[26] The cost-benefit calculation in this section suggests that the health benefits generated by LEZs outweigh the costs by 2.2 billion euros.

[26]This number, however, does not take into account the costs related to forgone working time for previously employed people. These costs according to a recent study by Weimar et al. (2003) average at 18,429 euros per employed patients.

General Conclusion

This dissertation examines three separate, yet interconnected dimensions of societal well-being. Chapter 1 poses the research question of whether schooling has a causal impact on attitudes toward immigration. Chapter 2 sets out to examine returns to internships completed during a course of studies and their role in a graduate's transition to the labor market. Chapter 3 homes in on the issue of air pollution in cities, and evaluates the effectiveness of a popular urban policy instrument in terms of pollution and health outcomes. To explore these research questions, each chapter uses several sets of microdata both from administrative sources and from surveys. To be able to tease out the causal effect, each study applies contemporary statistical and econometric approaches suited to the research question being considered. Chapters 1 and 2 apply an instrumental variable approach, while chapter 3 uses a difference-in-differences design. Each study tests the robustness of its findings to an array of sensitivity checks.

The first chapter shows that an additional year of schooling can have large civic returns, even in the absence of economic returns. The analysis finds that an additional year of schooling substantially reduces the likelihood of being very concerned about immigration in Germany. This study shows that the role of education in a society goes beyond the financial and labor market returns it generates. From the perspective of social cohesion, education plays an important role in shaping tolerant attitudes and breeding social trust. The findings in this chapter also show that educational policies may have an impact beyond the generation they directly affect. Understanding these intergenerational effects is important for quantifying the composite benefits of such policies, especially when weighing these benefits against costs. One limitation to chapter 1, and the related economic literature in general, is the incompleteness of information on parents and children captured simultaneously. Although this study uses rich survey data, it is still constrained by data limitations and a small sample size when it comes to shedding light on the transmission channels of parental education and children's civic beliefs. Exploring these channels could be a promising avenue for future research.

The second chapter of this thesis demonstrates the role of universities in facilitating the labor market transition of their graduates beyond academic teaching. This is the first study to show that completing firm internships has high returns in terms of wages and unemployment risk up to five to six years after graduation. The findings also indicate that certain groups may benefit more from such interventions. In the context of this empirical investigation, students graduating from areas of study traditionally less oriented toward specific job market skills, such as humanities, may see even larger returns to completed internships. These findings are of interest to individuals, educators and policymakers alike. Although the present evidence is based on an analysis of internships mandated by universities, such policies could easily be scaled toward centralized institutional bodies, as the experience of countries such as the UK and France shows. One drawback of the study is that the data we use includes no information on the type, duration and quality of the internships. Undoubtedly, designing an effective policy to steer the students in the right direction requires empirical knowledge on what constitutes effective internships. Further research would benefit from gathering and analyzing data that contains information on such specific characteristics.

The third chapter of this thesis investigates the effects of low emission zones (LEZs)—a popular air quality measure—on pollution and health. The analysis shows that LEZs reduce concentrations of two key pollutants (PM_{10} and NO_2) in treated areas and lead to health improvements for the individuals living inside these zones. These empirical findings have major policy implications. First, they contribute to the body of evidence regarding the effects of urban environmental policies and suggest that LEZs are an effective tool to address the issue of air quality. A back-of-the-envelope cost-benefit calculation also suggests that the health benefits outweigh the costs at the individual level. Second, this study shows that reductions in pollution, even starting at relatively low levels, may still generate health benefits, thus reinforcing the idea that there is no safe threshold for pollution. One important extension of this study could be an analysis of the distributional effects of LEZs. Since LEZs mainly target high-emission, and thus often older and cheaper cars, it is households from low socioeconomic backgrounds and small businesses that tend to bear the cost. The health data used in this paper, as is often the case, does not contain any information on the socioeconomic characteristics of the individuals. Additional work would be necessary to assess whether and to what extent LEZs are a regressive policy from a socioeconomic perspective. Lastly, the present analysis indicates that the improvements in air quality, while meaningful, are not large enough to reverse the existing health trajectories. If cities want to reduce air pollution further, stricter policy measures might be necessary.

To summarize, the three chapters in this thesis seek to study current and relevant dimensions of societal well-being. While the topics in this dissertation encompass important dimensions of quality of life, it cannot be claimed that the set of issues examined here is exhaustive. Determinants of societal well-being evolve and adapt to economic, social and political challenges. In fact, the overarching conclusion that can be drawn from the body of empirical evidence presented in this dissertation is that public policy has the scope and potential to achieve desirable outcomes across various dimensions of societal well-being with carefully designed measures.

List of Tables

List of Figures

Bibliography

Acemoglu, D. and J. Angrist (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual 15*(1), 9–59.

Akerlof, G. A. (1970). The market for lemons: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Alesina, A. and E. La Ferrara (2002). Who trusts others? *Journal of Public Economics 85*(2), 207–234.

Alexander, D. and H. Schwandt (2019). The impact of car pollution on infant and child health: Evidence from emissions cheating. Working Paper 12427, IZA.

Alternative für Deutschland (2016). Programm für Deutschland.

Altonji, J. G., L. B. Kahn, and J. D. Speer (2016). Cashier or consultant? Entry labor market conditions, field of study, and career success. *Journal of Labor Economics 34*(S1), S361–S401.

Andersen, Z. J., O. Raaschou-Nielsen, M. Ketzel, S. S. Jensen, M. Hvidberg, S. Loft, A. Tjønneland, K. Overvad, and M. Sørensen (2012). Diabetes incidence and long - term exposure to air pollution: A cohort study. *Diabetes Care 35*(1), 92–98.

Anderson, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association 18*(4), 1886–1927.

Angrist, J. D. and G. W. Imbens (1995). Two- stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association 90*(430), 431–442.

Angrist, J. D. and A. B. Krueger (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics 106*(4), 979–1014.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics 121*(1-2), 343–375.

Arulampalam, W. (2001). Is unemployment really scarring? Effects of unemployment experiences on wages. *The Economic Journal 111*(475), 585–606.

Ashenfelter, O. and A. Krueger (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review 84*(5), 1157–1173.

Auffhammer, M. and R. Kellogg (2011). Clearing the air? The effects of gasoline content regulation on air quality. *American Economic Review 101*(6), 2687–2722.

Avdeenko, A. and T. Siedler (2017). Intergenerational correlations of extreme right-wing party preferences and attitudes toward immigration. *The Scandinavian Journal of Economics 119*(3), 768–800.

Babadjouni, R. M., D. M. Hodis, R. Radwanski, R. Durazo, A. Patel, Q. Liu, and W. J. Mack (2017). Clinical effects of air pollution on the central nervous system; A review. *Journal of Clinical Neuroscience 43*, 16–24.

Bartl, W. and C. Korb (2009). Ost-West-Unterschiede bei der Studien- und Hochschulwahl: Ergebnisse der Studienanfängerbefragung an der Martin-Luther- Universität Halle-Wittenberg im Wintersemester 2008/09. *2009-1*, 48.

Beatty, T. K. M. and J. P. Shimshack (2011). School buses, diesel emissions, and respiratory health. *Journal of Health Economics 30*(5), 987–999.

Beck, J. E. and H. Halim (2008). Undergraduate internships in accounting: What and how do Singapore interns learn from experience? *Accounting Education 17*(2), 151–172.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3 ed.). Chicago: University of Chicago Press.

Becker, S. and F. Siebern-Thomas (2007). Schooling infrastructure, educational attainment and earnings. Mimeo.

Bishop, K. C., J. D. Ketcham, and N. V. Kuminoff (2018). Hazed and confused: The effect of air pollution on dementia. Working Paper 24970, National Bureau of Economic Research.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory 97*(2), 298–319.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *The American Economic Review 95*(1), 437–449.

Blom, A. G., C. Gathmann, and U. Krieger (2015). Setting up an online panel representative of the general population: The German Internet Panel. *Field Methods 27*(4), 391–408.

Bolton, J. L., N. C. Huff, S. H. Smith, S. N. Mason, W. M. Foster, R. L. Auten, and S. D. Bilbo (2013). Maternal stress and effects of prenatal air pollution on offspring mental health outcomes in mice. *Environmental health perspectives 121*(9), 1075–1082.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review 77*(4), 531–553.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education*, pp. 241–258. New York: Greenwood Press.

Box-Steffensmeier, J. M., S. De Boef, and T.-M. Lin (2004). The dynamics of the partisan gender gap. *American Political Science Review 98*(3), 515–528.

Braun, S. and M. Kvasnicka (2014). Immigration and structural change: Evidence from post-war Germany. *Journal of International Economics 93*(2), 253–269.

Braun, S. and T. O. Mahmoud (2014). The employment effects of immigration: Evidence from the mass arrival of German expellees in postwar Germany. *The Journal of Economic History 74*(1), 69–108.

Brook, R. D., B. Franklin, W. Cascio, Y. Hong, G. Howard, M. Lipsett, R. Luepker, M. Mittleman, J. Samet, S. C. Smith Jr, et al. (2004). Air pollution and cardiovascular disease: a statement for healthcare professionals from the expert panel on population and prevention science of the American Heart Association. *Circulation 109*(21), 2655–2671.

Brook, R. D., S. Rajagopalan, C. A. Pope, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters,

D. Siscovick, S. C. Smith, L. Whitsel, and J. D. Kaufman (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation 121*(21), 2331–2378.

Brooks, L., A. Cornelius, E. Greenfield, and R. Joseph (1995). The relation of career-related work or internship experiences to the career development of college seniors. *Journal of Vocational Behavior 46*(3), 332–349.

Bundesamt für Migration und Flüchtlinge (2016). Migration report 2015 - central conclusions.

Buoli, M., S. Grassi, A. Caldiroli, G. S. Carnevali, F. Mucci, S. Iodice, L. Cantone, L. Pergoli, and V. Bollati (2018). Is there a link between air pollution and mental disorders? *Environment international 118*, 154–168.

Busby, G. (2003). Tourism degree internships: A longitudinal study. *Journal of Vocational Education and Training 55*(3), 319–334.

Calderón-Garcidueñas, L. and R. Villarreal-Ríos (2017). Living close to heavy traffic roads, air pollution, and dementia. *The Lancet 389*(10070), 675–677.

Callanan, G. and C. Benzing (2004). Assessing the role of internships in the career-oriented employment of graduating college students. *Education + Training 46*(2), 82–89.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Canadian Undergraduate Survey Consortium (2004). Survey of first-year university students: 2004. *University of Victoria*.

Card, D. (1999). The causal effect of education on earnings. In *Handbook of Labor Economics*, Volume 3, Part A, pp. 1801–1863. Amsterdam: Elsevier.

Card, D., C. Dustmann, and I. Preston (2012). Immigration, wages, and compositional amenities. *Journal of the European Economic Association 10*(1), 78–119.

Carslaw, D. C. and G. Rhys-Tyler (2013). New insights from comprehensive on-road measurements of NOx, NO2 and NH3 from vehicle emission remote sensing in London, UK. *Atmospheric Environment 81*, 339–347.

Cesaroni, G., H. Boogaard, S. Jonkers, D. Porta, C. Badaloni, G. Cattani, F. Forastiere, and G. Hoek (2012). Health benefits of traffic-related air pollution reduction in different socioeconomic groups: the effect of low-emission zoning in Rome. *Occupational and Environmental Medicine 69*(2), 133–139.

Chau, T.-T. and K.-Y. Wang (2020). An association between air pollution and daily most frequently visits of eighteen outpatient diseases in an industrial city. *Scientific Reports 10*(1), 2321.

Chay, K. Y. and M. Greenstone (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics 118*(3), 1121–1167.

Coenders, M. and P. Scheepers (2003). The effect of education on nationalism and ethnic exclusionism: An international comparison. *Political Psychology 24*(2), 313–343.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology 94*, S95–S120.

Coneus, K. and C. K. Spiess (2012). Pollution exposure and child health: Evidence for infants and toddlers in Germany. *Journal of Health Economics 31*(1), 180–196.

Cook, S., R. S. Parker, and C. E. Pettijohn (2004). The perceptions of interns: A longitudinal case study. *Journal of Education for Business 79*(3), 179–185.

Currie, J. and E. Moretti (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics 118*(4), 1495–1532.

Currie, J. and M. Neidell (2005). Air pollution and infant health: What can we learn from California's recent experience? *The Quarterly Journal of Economics 120*(3), 1003–1030.

Currie, J., M. Neidell, and J. F. Schmieder (2009). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics 28*(3), 688–703.

Currie, J. and R. Walker (2011). Traffic congestion and infant health: Evidence from E-ZPass. *American Economic Journal: Applied Economics 3*(1), 65–90.

Cygan-Rehm, K. and M. Maeder (2013). The effect of education on fertility: Evidence from a compulsory schooling reform. *Labour Economics 25*, 35 – 48.

Davis, L. (2008). The effect of driving restrictions on air quality in Mexico City. *Journal of Political Economy 116*(1), 38–81.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics 88*(9–10), 1697–1720.

Delhey, J. and K. Newton (2005). Predicting cross-national levels of social trust: Global pattern or Nordic exceptionalism? *European Sociological Review 21*(4), 311–327.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review 109*(12), 4178–4219.

d'Hombres, B. and L. Nunziata (2016). Wish you were here? Quasi-experimental evidence on the effect of education on self-reported attitude toward immigrants. *European Economic Review 90*, 201–224.

Dickson, M., P. Gregg, and H. Robinson (2016). Early, late or never? When does parental education impact child outcomes? *The Economic Journal 126*(596), F184–F231.

Diegmann, V., F. Pfäfflin, G. Wiegand, and H. Wursthorn (2007). Massnahmen zur Reduzierung von Feinstaub und Stickstoffdioxid (Forschungsbericht 204 42 222; UBA-FB 000981). *Texte Umweltbundesamt* (22/07).

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies 79*(2), 645–677.

Durkheim, E. (2014). *The division of labor in society.* Simon and Schuster.

Dustmann, C. (2004). Parental background, secondary school track choice, and wages. *Oxford Economic Papers 56*(2), 209–230.

Dustmann, C. and I. Preston (2001). Attitudes to ethnic minorities, ethnic context and location decisions. *The Economic Journal 111*(470), 353–373.

Dustmann, C. and I. Preston (2007). Racial and economic factors in attitudes to immigration. *The B.E. Journal of Economic Analysis & Policy 7*(1).

Dustmann, C., U. Schönberg, and J. Stuhler (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives 30*(4), 31–56.

D'Amuri, F., G. I. P. Ottaviano, and G. Peri (2010). The labor market impact of immigration in Western Germany in the 1990s. *European Economic Review 54*(4), 550–570.

Economidou, C., D. Karamanis, A. Kechrinioti, and S. Xesfingi (2017). What shapes europeans' attitudes toward xeno-philia(/phobia)? *Munich Personal RePEc Archive, MPRA Paper No. 76511*.

Edlund, L. and R. Pande (2002). Why have women become left-wing? The political gender gap and the decline in marriage. *The Quarterly Journal of Economics 117*(3), 917–961.

Engelhardt, R. and K. Jahn (1964). *Politische Bildung im neunten Schuljahr: eine unterrichtspraktische Arbeitshilfe für die Volksschule*. Schule in Staat und Gesellschaft. Luchterhand.

Eurofound (2012). NEETs – young people not in employment, education or training: Characteristics, costs and policy responses in Europe. Technical report, Publications Office of the European Union, Luxembourg.

European Automobile Manufacturers Association (2021). Vehicles in use, Europe.

European Environment Agency (2016). Explaining road transport emissions: A non-technical guide.

Eurostat (2020). Unemployment by sex and age, annual data. `https://ec.europa.eu/eurostat/databrowser/view/UNE_RT_A__custom_492534/default/table`.

Eze, I. C., L. G. Hemkens, H. C. Bucher, B. Hoffmann, C. Schindler, N. Künzli, T. Schikowski, and N. M. Probst-Hensch (2015). Association between ambient air pollution and diabetes mellitus in europe and north america: Systematic review and meta-analysis. *Environmental Health Perspectives 123*(5), 381–389.

Fassmann, H. and R. Munz (1994). European East-West migration, 1945-1992. *The International Migration Review 28*(3), 520–538.

Fehr, E., U. Fischbacher, V. Rosenbladt, Bernhard, J. Schupp, and G. G. Wagner (2003). A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. Mimeo, Institute for Empirical Research in Economics, University of Zurich.

Fiordelisi, A., P. Piscitelli, B. Trimarco, E. Coscioni, G. Iaccarino, and D. Sorriento (2017). The mechanisms of air pollution and particulate matter in cardiovascular diseases. *Heart Failure Reviews 22*(3), 337–347.

Franchini, M. and P. M. Mannucci (2007). Short-term effects of air pollution on cardiovascular diseases: outcomes and mechanisms. *Journal of Thrombosis and Haemostasis 5*(11), 2169–2174.

Franz, W., J. Inkmann, W. Pohlmeier, and V. Zimmermann (1997). Young and out in Germany: On the youths' chances of labor market entrance in Germany. Working Paper 6212, National Bureau of Economic Research.

Fredriksson, P., L. Hensvik, and O. N. Skans (2018). Mismatch of talent: Evidence on match quality, entry wages, and job mobility. *American Economic Review 108*(11), 3303–38.

Freier, R., M. Schumann, and T. Siedler (2015). The earnings returns to graduating with honors — Evidence from law graduates. *Labour Economics 34*, 39–50.

Gallego, F., J.-P. Montero, and C. Salas (2013). The effect of transport policies on car use: Evidence from Latin American cities. *Journal of Public Economics 107*, 47–62.

Gangl, M. (2006). Scar effects of unemployment: An assessment of institutional complementarities. *American Sociological Review 71*(6), 986–1013.

Gault, J., J. Redington, and T. Schlager (2000). Undergraduate business internships and career success: Are they related? *Journal of Marketing Education 22*(1), 45–53.

Gehrsitz, M. (2017). The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management 83*, 121–144.

Gibson, M. and M. Carnovale (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics 89*, 62–73.

Glaeser, E. L. (2005). The political economy of hatred. *The Quarterly Journal of Economics 120*(1), 45–86.

Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany. *Journal of Labor Economics 30*(1), 175–213.

Gong, Y., S. Li, N. Sanders, and G. Shi (2019). The mortality impact of fine particulate matter in China. Working Paper.

Goodman-Bacon, A. (2018). Difference-in-Differences with variation in treatment timing. Working Paper 25018, National Bureau of Economic Research.

Goodwin, M. J. and O. Heath (2016). The 2016 referendum, Brexit and the left behind: An aggregate-level analysis of the result. *The Political Quarterly 87*(3), 323–332.

Graff Zivin, J. and M. Neidell (2013). Environment, health, and human capital. *Journal of Economic Literature 51*(3), 689–730.

Granovetter, M. S. (1995). *Getting a job: A study of contacts and careers* (2 ed.). University of Chicago Press.

Grave, B. S. and K. Görlitz (2012). Wage differentials by field of study – The case of German university graduates. *Education Economics 20*(3), 284–302.

Green, C. P., J. S. Heywood, and M. N. Paniagua (2020). Did the London congestion charge reduce pollution? *Regional Science and Urban Economics 84*, 103573.

Greenstone, M. and R. Hanna (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review 104*(10), 3038–3072.

Gregg, P. and E. Tominey (2005). The wage scar from male youth unemployment. *Labour Economics 12*(4), 487–509.

Hachmeister, C.-D., M. E. Harde, and M. F. Langer (2007). Einflussfaktoren der Studienentscheidung. Eine empirische Studie von CHE und EINSTIEG. Arbeitspapier 95, Centrum für Hochschulentwicklung, Gütersloh.

Hachmeister, C.-D. and M. Hennings (2007). Indikator im Blickpunkt: Kriterien der Hochschulwahl und Ranking-Nutzung. Auswertung aus dem CHE-Ranking. Technical report, Centrum für Hochschulentwicklung, Gütersloh.

Hainmueller, J. and M. J. Hiscox (2007). Educated preferences: Explaining attitudes toward immigration in Europe. *International Organization 61*(2), 399–442.

Hainmueller, J., M. J. Hiscox, and Y. Margalit (2015). Do concerns about labor market competition shape attitudes toward immigration? new evidence. *Journal of International Economics 97*(1), 193–207.

Hainmueller, J. and D. J. Hopkins (2014). Public attitudes toward immigration. *Annual Review of Political Science 17*(1), 225–249.

Halapuu, V., T. Paas, T. Tammaru, and A. Schütz (2013). Is institutional trust related to pro-immigrant attitudes? A pan-European evidence. *Eurasian Geography and Economics 54*(5-6), 572–593.

Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives 12*(2), 101–116.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings functions , rates of return and treatment effects : The mincer equation and beyond. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, pp. 307–458. Elsevier.

Heine, C., J. Schreiber, D. Sommer, and H. Spangenberg (2005). Studienanfänger in den Wintersemestern 2003/04 und 2004/05: Wege zum Studium, Studien- und Hochschulwahl, Situation bei Studienbeginn. Hochschulplanung 180, Hochschul-Informations-System, Hannover.

Heine, C., J. Willich, and H. Schneider (2009). Informationsverhalten und Hochschulwahl von Studienanfängern in West- und Ostdeutschland. Eine Sekundäranalyse der HIS Studienanfängerbefragung des Wintersemesters 2007/08. Projektbericht, Hochschul-Informations-System, Hannover.

Helliwell, J. F. and R. D. Putnam (1999). Education and social capital. Working Paper 7121, National Bureau of Economic Research.

Herreros, F. and H. Criado (2009). Social trust, social capital and perceptions of immigration. *Political Studies 57*(2), 337–355.

Holmlund, H., M. Lindahl, and E. Plug (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature 49*(3), 615–651.

Hong, Y.-C., J.-T. Lee, H. Kim, and H.-J. Kwon (2002). Air pollution: A new risk factor in ischemic stroke mortality. *Stroke 33*(9), 2165–2169.

Hoyt, J. E. and A. B. Brown (1999). Marketing UVSC: How prospective students view the college. mimeo.

Huang, J., H. Maassen van den Brink, and W. Groot (2009). A meta-analysis of the effect of education on social capital. *Economics of Education Review 28*(4), 454–464.

Huebener, M. and J. Marcus (2017). Compressing instruction time into fewer years of schooling and the impact on student performance. *Economics of Education Review 58*, 1–14.

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources 52*(4), 1060–1118.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Institut für Marktforschung GmbH (2014). Befragung zur Hochschulwahl: Online-Umfrage unter studierwilligen 16- bis 24-Jährigen im Auftrag der Länderübergreifenden Hochschulmarketingkampagne der ostdeutschen Länder.

Isen, A., M. Rossin-Slater, and W. R. Walker (2017). Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy 125*(3), 848–902.

Jacobson, L., R. LaLonde, and D. G. Sullivan (2005). Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics 125*(1-2), 271–304.

Janke, K. (2014). Air pollution, avoidance behaviour and children's respiratory health: Evidence from England. *Journal of Health Economics 38*, 23–42.

Jepsen, C., K. Troske, and P. Coomes (2014). The labor-market returns to community college degrees, diplomas, and certificates. *Journal of Labor Economics 32*(1), 95–121.

Jiang, W., M. Boltze, S. Groer, and D. Scheuvens (2017). Impacts of low emission zones in Germany on air pollution levels. *Transportation Research Procedia 25*, 3374–3386.

Kamhöfer, D. A. and H. Schmitz (2016). Reanalyzing zero returns to education in Germany. *Journal of Applied Econometrics 31*(5), 912–919.

Kemptner, D., H. Jürges, and S. Reinhold (2011). Changes in compulsory schooling and the causal effect of education on health : Evidence from Germany. *Journal of Health Economics 30*(2), 340–354.

Kemptner, D. and J. Marcus (2013). Spillover effects of maternal education on child's health and health behavior. *Review of Economics of the Household 11*(1), 29–52.

Klein, M. and F. Weiss (2011). Is forcing them worth the effort? Benefits of mandatory internships for graduates from diverse family backgrounds at labour market entry. *Studies in Higher Education 36*(8), 969–987.

Knack, S. and P. Keefer (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics 112*(4), 1251–1288.

Krawietz, M., P. Müßig-Trapp, and J. Willige (2006). HISBUS Blitzbefragung: Praktika im Studium. Kurzbericht Nr. 13, Hochschul-Informations-System, Hannover.

Lancee, B. and S. Pardos-Prado (2013). Group conflict theory in a longitudinal perspective : Analyzing the dynamic side of ethnic competition. *International Migration Review 47*(1), 106–131.

Lancee, B. and O. Sarrasin (2015). Educated preferences or selection effects? A longitudinal analysis of the impact of educational attainment on attitudes towards immigrants. *European Sociological Review 31*(4), 490–451.

Lange, F., K. Kroft, and M. J. Notowidigdo (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics 128*(3), 1123–1167.

Law, M. R., J. Morris, and N. J. Wald (1997). Environmental tobacco smoke exposure and ischaemic heart disease: An evaluation of the evidence. *Bmj 315*(7114), 973–980.

Leal, J., R. Luengo-Fernández, A. Gray, S. Petersen, and M. Rayner (2006). Economic burden of cardiovascular diseases in the enlarged European Union. *European Heart Journal 27*(13), 1610–1619.

Lee, B.-J., B. Kim, and K. Lee (2014). Air pollution exposure and cardiovascular disease. *Toxicological Research 30*(2), 71–75.

Lee, K. K., M. R. Miller, and A. S. Shah (2018). Air pollution and stroke. *Journal of Stroke 20*(1), 2–11.

Leiva, G, M. A., D. A. Santibañez, S. Ibarra E, P. Matus C, and R. Seguel (2013). A five-year study of particulate matter (PM2.5) and cerebrovascular diseases. *Environmental Pollution (Barking, Essex: 1987) 181*, 1–6.

Leschinsky, A. and P. M. Roeder (1980). Didaktik und Unterricht in der Sekundarschule I seit 1950 - Entwicklung der Rahmenbedingungen. In J. Baumert, A. Leschinsky, J. Naumann, J. Raschert, and P. Siewert (Eds.), *Bildung in der Bundesrepublik Deutschland - Daten und Analysen, Entwicklungen seit 1950*, Volume 1, pp. 283–392. Stuttgart: Klett-Cotta.

Leuze, K. and S. Strauß (2009). Lohnungleichheiten zwischen Akademikerinnen und Akademikern: Der Einfluss von fachlicher Spezialisierung, frauendominierten Fächern und beruflicher Segregation. *Zeitschrift für Soziologie 38*(4), 262–281.

Linn, W. S., Y. Szlachcic, H. Gong, P. L. Kinney, and K. T. Berhane (2000). Air pollution and daily hospital admissions in metropolitan Los Angeles. *Environmental Health Perspectives 108*(5), 427–434.

Lisabeth, L. D., J. D. Escobar, J. T. Dvonch, B. N. Sánchez, J. J. Majersik, D. L. Brown, M. A. Smith, and L. B. Morgenstern (2008). Ambient air pollution and risk for ischemic stroke and transient ischemic attack. *Annals of Neurology 64*(1), 53–59.

Lleras-Muney, A. (2010). The needs of the army: Using compulsory relocation in the military to estimate the effect of air pollutants on children's health. *Journal of Human Resources 45*(3), 549–590.

Lochner, L. (2011). Non-production benefits of education: Crime, health, and good citizenship. Working Paper 16722, National Bureau of Economic Research.

Lundborg, P., A. Nilsson, and D.-O. Rooth (2014). Parental education and offspring outcomes: Evidence from the Swedish compulsory school reform. *American Economic Journal: Applied Economics 6*(1), 253–278.

Machin, S. and P. A. Puhani (2003). Subject of degree and the gender wage differential: Evidence from the UK and Germany. *Economics Letters 79*(3), 393–400.

Maheswaran, R., T. Pearson, N. C. Smeeton, S. D. Beevers, M. J. Campbell, and C. D. Wolfe (2012). Outdoor air pollution and incidence of ischemic and hemorrhagic stroke: A small-area level ecological study. *Stroke 43*(1), 22–27.

Malina, C. and F. Scheffler (2015). The impact of Low Emission Zones on particulate matter concentration and public health. *Transportation Research Part A: Policy and Practice 77*, 372–385.

Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *Review of Economics and Statistics 88*(3), 510–530.

McHenry, P. (2015). Immigration and the human capital of natives. *Journal of Human Resources 50*(1), 34–71.

Mehring, M., E. Donnachie, R. Mutschler, F. Hofmann, M. Keller, and A. Schneider (2013). Disease management programs for patients with asthma in Germany: A longitudinal population-based study. *Respiratory Care 58*(7), 1170–1177.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). Does education improve citizenship? Evidence from the United States and the United Kingdom. *Journal of Public Economics 88*(9–10), 1667–1695.

Mills, N. L., K. Donaldson, P. W. Hadoke, N. A. Boon, W. MacNee, F. R. Cassee, T. Sandström, A. Blomberg, and D. E. Newby (2009). Adverse cardiovascular effects of air pollution. *Nature Clinical Practice Cardiovascular Medicine 6*(1), 36–44.

Mincer, J. (1974). *Schooling, experience, and earnings.* New York: Columbia University Press.

Mocan, N. and C. Raschke (2016). Economic well-being and anti-semitic, xenophobic, and racist attitudes in Germany. *European Journal of Law and Economics 41*(1), 1–63.

Moretti, E. and M. Neidell (2011). Pollution, health, and avoidance behavior: Evidence from the ports of Los Angeles. *The Journal of Human Resources 46*(1), 154–175.

Morfeld, P., D. A. Groneberg, and M. F. Spallek (2014). Effectiveness of Low Emission Zones: Large scale analysis of changes in environmental NO2, NO and NOx concentrations in 17 German cities. *PLoS ONE 9*(8), e102999.

Mroz, T. A. and T. H. Savage (2006). The long-term effects of youth unemployment. *The Journal of Human Resources 41*(2), 259–293.

Neidell, M. (2009). Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations. *Journal of Human Resources 44*(2), 450–478.

Neidell, M. J. (2004). Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma. *Journal of Health Economics 23*(6), 1209–1236.

Nicholls, A. J. (1978). The British impact on German education: a triumph for commonsense or missed opportunity? *Oxford Review of Education 4*(2), 125–129.

Nordrhein-Westfalen (1962). Landtagsdrucksache. *4. Wahlperiode, Band 5, Nr. 696.*

Nunley, J. M., A. Pugh, N. Romero, and R. A. Seals (2017). The effects of unemployment and underemployment on employment opportunities: Results from a correspondence audit of the labor market for college graduates. *ILR Review 70*(3), 642–669.

Nunley, J. M., A. Pugh, N. Romero, and R. A. Seals Jr. (2016). College major, internship experience, and employment opportunities: Estimates from a résumé audit. *Labour Economics 38*, 37–46.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *The American Economic Review 96*(1), 152–175.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics 91*(11-12), 2213–2229.

Oreopoulos, P. and K. G. Salvanes (2011). Priceless: The nonpecuniary benefits of schooling. *The Journal of Economic Perspectives 25*(1), 159–184.

O'Rourke, K. H. and R. Sinnott (2006). The determinants of individual attitudes towards immigration. *European Journal of Political Economy 22*(4), 838–861.

Parey, M. and F. Waldinger (2011). Studying abroad and the effect on international labour market mobility: Evidence from the introduction of ERASMUS. *Economic Journal 121*(551), 194–222.

Pei, Z., J.-S. Pischke, and H. Schwandt (2019). Poorly measured confounders are more useful on the left than on the right. *Journal of Business & Economic Statistics 37*(2), 205–216.

Pelkonen, P. (2012). Length of compulsory education and voter turnout—Evidence from a staged reform. *Public Choice 150*(1-2), 51–75.

Pestel, N. and F. Wozny (2019). Low emission zones for better health: Evidence from German hospitals. Working Paper 12545, IZA.

Peters, R., N. Ee, J. Peters, A. Booth, I. Mudway, and K. J. Anstey (2019). Air pollution and dementia: A systematic review. *Journal of Alzheimer's Disease 70*(s1), S145–S163.

Petzold, H.-J. (1981). *Schulzeitverlängerung: Parkplatz oder Bildungschance? Die Funktion des 9. und 10. Schuljahres.* Bensheim: päd extra Buchverlag.

Piopiunik, M. (2014). Intergenerational transmission of education and mediating channels: Evidence from a compulsory schooling reform in Germany. *Scandinavian Journal of Economics 116*(3), 878–907.

Pischke, J.-S. (2007). The impact of length of the school year on student performance and earnings: Evidence from the German short school years. *The Economic Journal 117*(523), 1216–1242.

Pischke, J.-S. and T. von Wachter (2005). Zero returns to compulsory schooling in Germany: Evidence and interpretation. Working Paper 11414, National Bureau of Economic Research.

Pischke, J.-S. and T. von Wachter (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *Review of Economics and Statistics 90*(3), 592–598.

Plug, E. (2004). Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *The American Economic Review 94*(1), 358–368.

Pope, I. C. (2000). Epidemiology of fine particulate air pollution and human health: Biologic mechanisms and who's at risk? *Environmental health perspectives 108*(suppl 4), 713–723.

Poutvaara, P. and M. F. Steinhardt (2018). Bitterness in life and attitudes towards immigration. *European Journal of Political Economy 55*, 471–490.

Power, M. C., S. D. Adar, J. D. Yanosky, and J. Weuve (2016). Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. *NeuroToxicology 56*, 235–253.

Pryor, J. H., K. Eagan, L. P. Blake, S. Hurtado, J. Berdan, and M. H. Case (2012). The American freshman: National norms fall 2012. Technical report, Higher Education Research Institute, Graduate School of Education & Information Studies, University of California, Los Angeles.

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy 6*(1), 65–78.

Raaum, O. and K. Røed (2006). Do business cycle conditions at the time of labor market entry affect future employment prospects? *The Review of Economics and Statistics 88*(2), 193–210.

Rehn, T., G. Brandt, G. Fabian, and K. Briedis (2011). Hochschulabschlüsse im Umbruch: Studium und Übergang von Absolventinnen und Absolventen reformierter und traditioneller Studiengänge des Jahrgangs 2009. HIS: Forum Hochschule 17, Hochschul-Informations-System, Hannover.

Reimer, D. and J. Schröder (2006). Tracing the gender wage gap: Income differences between male and female university graduates in Germany. *Journal for Labour Market Research 39*(2), 235–253.

Restrepo, B. J. and M. Rieger (2016). Trans fat and cardiovascular disease mortality: Evidence from bans in restaurants in New York. *Journal of Health Economics 45*, 176–196.

Richards, E. W. (1984). Undergraduate preparation and early career outcomes: A study of recent college graduates. *Journal of Vocational Behavior 24*(3), 279–304.

Riphahn, R. T. (2003). Cohort effects in the educational attainment of second generation immigrants in Germany: An analysis of census data. *Journal of Population Economics 16*(4), 711–737.

Ruhm, C. J. (1991). Are workers permanently scarred by job displacements? *The American economic review 81*(1), 319–324.

RWI-GEO-GRID (2019). RWI-GEO-GRID: Socio-economic data on grid level-scientific use file (wave 8). version: 1. Leibniz Institute for Economic Research. `http://doi.org/10.7807/microm:suf:V8`.

RWI-GEO-RED. RWI Real Estate Data (scientific use file)- apartments for rent. `http://fdz.rwiessen.de/doi-detail/id-107807immoredwmsufv1.html`.

Rydgren, J. and P. Ruth (2011). Voting for the radical right in Swedish municipalities: Social marginality and ethnic competition? *Scandinavian Political Studies 34*(3), 202–225.

Sanders, N. J. and C. Stoecker (2015). Where have all the young men gone? Using sex ratios to measure fetal death rates. *Journal of Health Economics 41*, 30–45.

Saniter, N. (2012). Estimating heterogeneous returns to education in Germany via conditional heteroskedasticity. IZA Discussion Paper 6813.

Sarcletti, A. (2009). Die Bedeutung von Praktika und studentischen Erwerbstätigkeiten für den Berufseinstieg: Dissertation. Studien zur Hochschulforschung 77, Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung, München.

Scheve, K. F. and M. J. Slaughter (2001). Labor market competition and individual preferences over immigration policy. *Review of Economics and Statistics 83*(1), 133–145.

Schiefer, D. and J. van der Noll (2017). The essentials of social cohesion: A literature review. *Social Indicators Research 132*(2), 579–603.

Schlenker, W. and W. R. Walker (2016). Airports, air pollution, and contemporaneous health. *Review of Economic Studies 83*(2), 768–809.

Schnedler, W. (2004). *The value of signals in hidden action models: Concepts, application, and empirical evidence.* Contributions to economics. Heidelberg: Physica-Verlag.

Science for Environment Policy (2021). What are the health costs of environmental pollution? Future brief 21. Brief produced for the European Commission DG Environment by the Science Communication Unit.

Shoenfelt, E. L., N. J. Stone, and J. L. Kottke (2013). Internships: An established mechanism for increasing employability. *Industrial and Organizational Psychology 6*(1), 24–27.

Siedler, T. (2010). Schooling and citizenship in a young democracy: Evidence from postwar Germany. *Scandinavian Journal of Economics 112*(2), 315–338.

Simeonova, E., J. Currie, P. Nilsson, and R. Walker (2018). Congestion pricing, air pollution and children's health. Working Paper 24410, National Bureau of Economic Research.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics 87*(3), 355–374.

Spiess, K. C. and K. Wrohlich (2010). Does distance determine who attends a university in Germany? *Economics of Education Review 29*(3), 470–479.

Stafoggia, M., G. Cesaroni, A. Peters, Z. J. Andersen, C. Badaloni, R. Beelen, B. Caracciolo, J. Cyrys, U. de Faire, K. de Hoogh, K. T. Eriksen, L. Fratiglioni, C. Galassi, B. Gigante, A. S. Havulinna, F. Hennig, A. Hilding, G. Hoek, B. Hoffmann, D. Houthuijs, M. Korek, T. Lanki, K. Leander, P. K. Magnusson, C. Meisinger, E. Migliore, K. Overvad, C.-G. Ostenson, N. L. Pedersen, J. Pekkanen, J. Penell, G. Pershagen, N. Pundt, A. Pyko, O. Raaschou-Nielsen, A. Ranzi, F. Ricceri, C. Sacerdote, W. J. R. Swart, A. W. Turunen, P. Vineis, C. Weimar, G. Weinmayr, K. Wolf, B. Brunekreef, and F. Forastiere (2014). Long-term exposure to ambient air pollution and incidence of cerebrovascular events: Results from 11 European cohorts within the ESCAPE project. *Environmental Health Perspectives 122*(9), 919–925.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Statistisches Bundesamt (2014). Studierende an Hochschulen, Fächersystematik: Wintersemester 2013/2014. Fachserie 11 Reihe 4.1, Statistisches Bundesamt, Wiesbaden.

Statistisches Bundesamt (2015). Krankheitsklassen und Alter in Euro je Einwohner der jeweiligen Altersgruppe.

Statistisches Bundesamt (2017). Herz-Kreislauf-Erkrankungen verursachen die höchsten Kosten. Pressemitteilung vom 29. September 2017 - 347/17.

Stephens, M. and D.-Y. Yang (2014). Compulsory education and the benefits of schooling. *The American Economic Review 104*(6), 1777–1792.

Steppuhn, H., U. Langen, S. Mueters, S. Dahm, H. Knopf, T. Keil, and C. Scheidt-Nave (2016). Asthma management practices in adults–findings from the German Health Update (GEDA) 2010 and the German National Health Interview and Examination Survey (DEGS1) 2008–2011. *Journal of Asthma 53*(1), 50–61.

Stiglitz, J. E. (1975). The theory of "screening," education, and the distribution of income. *65*(3), 283–300.

Stiglitz, J. E., A. Sen, J.-P. Fitoussi, et al. (2009). Report by the commission on the measurement of economic performance and social progress.

Sun, C., S. Zheng, and R. Wang (2014). Restricting driving for better traffic and clearer skies: Did it work in Beijing? *Transport Policy 32*, 34–41.

Taylor, M. S. (1988). Effects of college internships on individual participants. *Journal of Applied Psychology 73*(3), 393–401.

Teichler, U. (2011). Bologna — Motor or stumbling block for the mobility and employability of graduates? In H. Schomburg and U. Teichler (Eds.), *Employability and mobility of bachelor graduates in Europe*, pp. 3–41. Rotterdam: SensePublishers.

Tent, J. F. (1982). Mission on the rhine: American educational policy in postwar germany, 1945–1949. *History of Education Quarterly 22*(3), 255–276.

Thiering, E. and J. Heinrich (2015). Epidemiology of air pollution and diabetes. *Trends in Endocrinology & Metabolism 26*(7), 384–394.

Tsai, S.-S., W. B. Goggins, H.-F. Chiu, and C.-Y. Yang (2003). Evidence for an association between air pollution and daily stroke admissions in Kaohsiung, Taiwan. *Stroke 34*(11), 2612–2616.

Umweltbundesamt (2018). Air quality 2017. Preliminary evaluation. Technical report.

UN (2018). World urbanization prospects 2018. United Nations Department for Economic and Social Affairs.

Van Heerden, S. and D. Ruedin (2019). How attitudes towards immigrants are shaped by residential context: The role of ethnic diversity dynamics and immigrant visibility. *Urban Studies 56*(2), 317–334.

Viard, V. B. and S. Fu (2015). The effect of Beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics 125*, 98–115.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP): Scope, evolution and enhancements. SOEPpapers on Multidisciplinary Panel Data Research, DIW Berlin, The German Socio-Economic Panel (SOEP).

Weimar, C., C. Weber, M. Wagner, O. Busse, R. L. Haberl, K. W. Lauterbach, H. C. Diener, and German Stroke Data Bank Collaborators (2003). Management patterns and health care use after intracerebral hemorrhage. a cost-of-illness study from a societal perspective in Germany. *Cerebrovascular Diseases 15*(1-2), 29–36.

Wellenius, Gregory, A., J. Schwartz, and M. A. Mittleman (2005). Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries. *Stroke 36*(12), 2549–2553.

Wolff, H. (2014). Keep your clunker in the suburb: Low emission zones and adoption of green vehicles. *The Economic Journal 124*(578), F481–F512.

Wolff, H. and L. Perry (2010). Policy monitor: Trends in clean air legislation in Europe: Particulate matter and Low Emission Zones. *Review of Environmental Economics and Policy 4*(2), 293–308.

Wolter, A. and U. Banscherus (2012). Praxisbezug und Beschäftigungsfähigkeit im Bologna-Prozess — A Never Ending Story? In W. Schubarth, K. Speck, A. Seidel, C. Gottmann, C. Kamm, and M. Krohn (Eds.), *Studium nach Bologna: Praxisbezüge stärken?*, pp. 21–36. Wiesbaden: Springer Fachmedien.

Worbs, S. (2003). The second generation in Germany: Between school and labor market. *International Migration Review 37*(4), 1011–1038.

Xue, T., T. Zhu, Y. Zheng, and Q. Zhang (2019). Declines in mental health associated with air pollution and temperature variability in China. *Nature communications 10*(1), 1–8.

Zak, P. J. and S. Knack (2001). Trust and growth. *The Economic Journal 111*(470), 295–321.

Zhang, P., G. Dong, B. Sun, L. Zhang, X. Chen, N. Ma, F. Yu, H. Guo, H. Huang, Y. L. Lee, et al. (2011). Long-term exposure to ambient air pollution and mortality due to cardiovascular disease and cerebrovascular disease in Shenyang, China. *PloS one 6*(6).

Zhang, W., C. Y. C. Lin Lawell, and V. I. Umanskaya (2017). The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence. *Journal of Environmental Economics and Management 82*, 181–220.

Declarations

Declaration

I hereby declare that I, Shushanik Margaryan, have not received any commercial consultation on my doctoral thesis. This thesis has not been accepted as part of any previous doctoral procedure or graded as insufficient.

Hamburg, 25 February 2021

Shushanik Margaryan

Affidavit

I, hereby declare under oath that I wrote the dissertation titled "Empirical Analyses of Societal Challenges: Social Cohesion, Labor Market Transition and Population Health" myself and in case of cooperation with other researchers pursuant to the enclosed statements in accordance with Section 6 subsection 3 of the Doctoral Degree Regulations of the Faculty of Business, Economics and Social Sciences dated 18 January 2017. I have used no aids other than those indicated: Excel, Latex, Stata, QGIS, and the literature indicated in the bibliography.

Hamburg, 25 February 2021

Shushanik Margaryan

Personal declaration for dissertation by publication

Chapter 1

Chapter 1 is written in co-authorship with Thomas Siedler and Annemarie Paul. The Chapter is accepted for published under the title "Does immigration affect attitudes towards migration: Evidence from Germany" in the *Journal of Human Resources*. Shushanik Margaryan's personal contribution breaks down as:

- Concept/planning: 50%
- Implementation: 80%
- Creation of manuscript: 50%

Chapter 2

Chapter 2 is written in co-authorship with Nils Saniter, Mathias Schumann and Thomas Siedler. The Chapter is accepted for publication under the title "Do Internships Pay Off? The Effects of Student Internships on Earnings" in the *Journal of Human Resources*. Shushanik Margaryan's personal contribution breaks down as:

- Concept/planning: 25%
- Implementation: 25%
- Creation of manuscript: 25%

Chapter 3

Shushanik Margaryan's personal contribution to Chapter 3 is 100% in conception, planning, implementation and creation of the manuscript. The Chapter is published under the title "Low Emission Zones and Population Health" in the *Journal of Health Economics*.

Hamburg, 25 February 2021

Shushanik Margaryan