Empirical Essays on Human Capital Investments in Health and Education

Anastasia Driva

Munich 2016

Empirical Essays on Human Capital Investments in Health and Education

Anastasia Driva

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae
(Dr. oec. publ.)
an der Ludwig–Maximilians–Universität München

2016

vorgelegt von
Anastasia Driva

Referent: Prof. Dr. Joachim Winter

Korreferent: Prof. Davide Cantoni, Ph.D.

Promotionsabschlussberatung: 10. Mai 2017

Tag der mündlichen Prüfung: 4. Mai 2017

Namen der Berichterstatter: Joachim Winter, Davide Cantoni, Amelie Wuppermann

Acknowledgements

First and foremost, I would like to thank my supervisor and first year mentor, Joachim Winter, for his support and valuable advice. He always had an open door for me and I highly appreciate that. I would further like to thank Davide Cantoni and Amelie Wuppermann for their insightful comments and for agreeing to serve on my committee.

In addition, I am grateful to my co-authors in different projects, Stefan Bauernschuster, Erik Hornung and Melanie Lührmann for a great cooperation and for their positive attitude. I have very much enjoyed working with them and I have definitely learnt a lot.

I would also like to thank Till von Wachter and the UCLA graduates for making everything possible so that I have an inspiring and pleasant time at the University of California, Los Angeles. My visit, although short, was definitely one of the most interesting and memorable experiences of my Ph.D. !

I greatly appreciate the generous support from the Elite Network of Bavaria and the Evidence-Based Economics program, in particular for providing access to a breadth of research opportunities. I am grateful for having taken part in the Lindau Nobel Laureate Meetings in Economic Sciences, for having presented in numerous conferences and for having attended intellectually stimulating guest courses and seminars. To my colleagues at the Munich Graduate School of Economics, it has been a real pleasure meeting them all and I am happy for the moments we shared together.

Finally, I am indebted to my family and good friends for their immeasurable support in good and more difficult times. This dissertation would not have been the same without them.

Anastasia Driva, December 2016

Table of Contents

List of Figures

List of Tables

Preface

Health and education have been described as the "twin pillars for assuring the well-being of individuals" by the former Secretary General of the United Nations, Kofi Annan (Khan, 2002). According to data from the Worldbank, OECD member-countries spend on average 12.4% and 4.9% of total GDP on health and education respectively. The figures are lower for low and middle income countries but still sizable (World Bank, 2016b). The interaction of the two is of high relevance not only for governments and policymakers but also for economists. Long-run growth theory has emphasized the instrumental role of human capital investments, a subset of which is health and education, for economic growth (Preston, 1975; Mincer, 1981; Costa and Steckel, 1997; Bloom et al., 2001).

A necessary condition for policymakers and economists to foster health and education is first to quantify the two notions and understand stylized facts and second to design policies to improve them both in the short- and in the long-term. From an empirical point of view, there has been a plethora of studies and interventions targeting improvements in health and education outcomes in developed and developing countries (Aizer, 2004; Lleras-Muney, 2005; Heckman et al., 2013; Attanasio, 2015). Yet, there is still no consensus with respect to what type of interventions work best under specific settings. To shed light on the importance of reforms and interventions in health and education, I look into three different settings in Germany, where I combine theoretical foundations with statistical tools and data to empirically investigate questions of high policy-interest that have been neglected by the literature so far. In particular, this dissertation consists of three separate chapters that assess the role of introducing public health insurance on mortality, the relationship between financial knowledge and stereotypes and the role of expanding childcare for health outcomes.

The theoretical foundations underlying the importance of interventions in health and education relate to earlier work by Becker (1975), Preston (1975) and Mincer (1981) as well as Carneiro and Heckman (2003), Cunha et al. (2006) and Bleakley (2007) more recently. In a seminal paper, Preston (1975) was the first to document that a large part of the increase in life expectancy observed during the 20th century can be attributed to public health interventions such as vaccination, technology and public health infrastructure as well as better education. Becker (1975) and Bleakley (2007) focus on what is known as the quantity-quality tradeoff, i.e. the idea that increasing expenditures on the quality of individual children (e.g. via education) raise the costs of fertility and hence the decision to have more children. Building on the work of Becker (1975), Carneiro and Heckman (2003) and Cunha et al. (2006) highlight the importance of skill formation for later life outcomes. In a theoretical framework, the authors provide an explanation of how health and education interventions could yield higher returns when made early in life. Skill formation is a life-long process where each stage corresponds to a period in the life cycle of a child. Inputs at each stage produce outputs, i.e. skills, at that stage. Some inputs may be more productive at certain stages and childhood is an example of such a critical period. Human capital investments during that period could act either as springboards for later accumulation of skills or as facilitators for higher skill attainment later in life.

Empirical findings confirm the aforementioned hypotheses both when interventions take place early in life or at a later stage. For instance, adolescence is a crucial period during which many changes in a teenager's life take place. From a health perspective, teenagers form habits (e.g. dietary habits, smoking, drugs, alcohol) while from an education point of view, interventions could encourage the youth to invest more in improving their knowledge on financial matters, a correlate of wealth. Lührmann et al. (2015b) examine the impact of a financial education program on adolescents in Germany to find significant improvements in their financial knowledge and their interest in financial matters. Furthermore, it is an established fact that educational choices and labour market outcomes are strongly correlated. Higher secondary school tracks are associated with better labour market opportunities, in particular for disadvantaged individuals (Dustmann, 2004). Preschool interventions are also shown to yield better health and education outcomes as well as higher earnings in the long-run (Cunha et al., 2006; Ludwig and Miller, 2007; Havnes and Mogstad, 2011b). An evaluation of the Head Start program by Ludwig and Miller (2007) documents that there is a large drop in mortality rates for children who were treated. Havnes and Mogstad (2011b) further show that subsidized childcare has a positive and significant effect on children's educational attainment.

Chapter 1 is a joint project with Stefan Bauernschuster and Erik Hornung which looks into the impact of compulsory health insurance on mortality. In particular, we show that the world's first compulsory health insurance scheme, as established by Chancellor Bismarck in 1884, causally reduced mortality of blue-collar workers and the effects are large and statistically significant. The analysis is motivated by the role that public health investments played on reducing population mortality during the demographic transition. Understanding the impact of public interventions for demographic change and economic growth is crucial for the design of effective policies. A large strand of the literature provides evidence on how public health initiatives such as improvements in water supply or the eradication of hookworm and malaria improved life expectancy (Cutler and Miller, 2005; Bleakley, 2007; Costa, 2015). In another case, Wüst (2012) finds that Danish infant mortality was reduced due to home visiting programs which helped towards diffusing knowledge on nutrition as well as highlighting the positive health effects of breastfeeding. According to findings by Bütikofer et al. (2015), Norwegian healthcare centers also contributed to better long-term economic outcomes by providing home visits during the first year of life.

Using novel and unique data, our analysis contributes to this literature by answering an open question that had long been neglected, namely the role that compulsory public health insurance played for population's health during the demographic transition. While recent expansions of compulsory health insurance have been highly effective in increasing access to healthcare and reducing mortality (see Finkelstein, 2007; Card et al., 2008), there are key differences between these 20th century expansions and our setting. Current health insurance schemes work in an environment of chronic diseases and have the ability to provide healthcare by medical treatment. In contrast, we investigate the impact of compulsory health insurance in a setting of infectious diseases where there is no effective medical treatment available. This approach speaks to the situation in many developing countries today, where universal healthcare is absent while infectious diseases are responsible for a large part of premature deaths (World Health Organization, 2016b).

Our empirical approach exploits differences in insurance eligibility across occupational groups. To our knowledge, we are the first to provide a causal interpretation of Bismarck's Health Insurance on mortality by disentangling it from other channels of public health investments such as waterworks and sewerage. To quantify the effect of the reform, we draw on unique and novel administrative data from Prussia. The data goes back to the 19th century and includes rich district-level information on the number of deceased by occupational groups, census and population variables as well as city-level information on public infrastructure, namely waterworks and sewerage. We combine Prussian administrative data with time-series, longitudinal data from other countries to undertake an international comparison between mortality in Prussia and mortality in other countries.

We find that after the introduction of Bismarck's Health Insurance, mortality for blue-collar workers decreases significantly. Additional evidence suggests that a large part of the reduction in mortality is driven by a decline in deaths from airborne infectious diseases. A potential channel for this finding is the access to physicians that the insured had and especially the information passed on by doctors with respect to hygiene rather than the effectiveness of treatment itself. These findings are supportive of the hypothesis that the health insurance scheme provided families at the lower end of the income distribution with access to knowledge related to hygiene.

Chapter 2 of my dissertation — a joint project with Melanie Lührmann and Joachim Winter — moves away from public health interventions and looks into financial literacy and gender stereotypes among teenagers. Jappelli and Padula (2013) point out that financial knowledge can be accumulated from an early point in life, similar to other forms of human capital. The authors argue that the decision to invest in financial literacy may yield high future returns in the form of savvier investment decisions but it also involves short-term effort and time cost. This tradeoff between short-term costs and long-term benefits is not always salient to individuals. The lack of integration of financial education courses in school curricula makes this tradeoff more tenuous. Lührmann et al. (2015b) for instance find that financial literacy training programs significantly increase teenagers' interest in financial matters. Another strand of the literature has shown that the likelihood of investing in the stock market is strongly correlated with cognitive abilities (Christelis et al., 2010). Finally, researchers consistently find a gender gap with respect to financial knowledge across countries (Lusardi and Mitchell, 2008; Lusardi et al., 2010).

In this study, we explore the relationship between gender differences and stereotypes in financial literacy among teenagers. Financial literacy is an important correlate of financial wealth and a reliable predictor of stock market participation, savings and retirement planning (Lusardi and Mitchell, 2008; van Rooij et al., 2011). The determinants of the gender gap in financial literacy are largely unknown. Most studies that look into the gender gap in financial literacy focus on adults. Lührmann et al. (2015b) is an exception to this as they document that the gender gap in financial literacy already exists at younger ages, i.e. among teenagers between 13 and 15 years old. We go one step further and combine this literature with recent findings showing that stereotypes can explain gender gaps in various domains (Bordalo et al., 2016).

In our analysis, we use a sample of 418 high-school students which are recruited from thirty classes, across thirteen schools, in three German cities. We start off by confirming the existence of the financial knowledge gap between boys and girls in their teens. These gender differences in financial literacy may be related to self-confidence, gender-specific risk attitudes, or numeracy. In a second step, we report an association between financial knowledge and gender stereotypes related to household finance. Using a stereotype index that measures strength of finance-related beliefs, we find that males' beliefs are more biased towards their own gender. On the contrary, females do not exhibit such self-affirmative beliefs. In other words, both genders believe that males are more competent with respect to either questions related to interest and ability in finance or to whether men are more likely to be concerned with finances at home and at work. The policy relevance of our findings lies in that they suggest possible ways to improve financial education programs targeted at younger individuals. Their effectiveness for females might be increased by addressing stereotypes directly. To our knowledge, we are the first ones to provide evidence that builds a bridge between financial literacy and gender stereotypes among teenagers.

In the third chapter of this dissertation, I study the effect of being eligible for earlier childcare entry on long-term health outcomes. My analysis is motivated by the seminal work of Carneiro and Heckman (2003) and Cunha et al. (2006) in which the authors highlight the important role that early education investments play for long-term outcomes. Inadequate investment in either cognitive or non-cognitive skills, for example, could account for inequalities in schooling, health and other socioeconomic dimensions later in life. From a policy perspective, the Worldbank estimates that "every dollar spent on preschool education earns between 6 and 17 dollars of public benefits, in the form of a healthier and more productive workforce" (World Bank, 2016a). Therefore, looking at the relationship between childcare and health is of increasing relevance both for economists and policymakers.

To the best of my knowledge, I am the first to investigate the long-term implications of childcare expansion on health by exploiting this specific natural experiment in Germany. In particular, I look into the nationwide introduction of a legal claim to a place in kindergarten, first implemented in 1996. This mandated municipalities to provide childcare to children as early as from the age of three. In addition, this entitlement made entry to formal childcare conditional on a date-of-birth cutoff rule. Based on this rule — children who were above 36 months old in August or September of the kindergarten year — were allowed to enter kindergarten in this year at the age of three. Before this rule, most children entered childcare either at the age of four or five. In the years that followed the reform, there was a substantial increase in the number of children attending formal childcare at the age of three. Based on findings from the early childhood interventions literature (Carneiro and Heckman, 2003; Cunha et al., 2006), I hypothesize that being eligible for childcare earlier is associated with better long-term behavioural health outcomes such as reduced smoking and better dietary habits. To investigate this hypothesis, I use data from the German Socio-Economic Panel study (SOEP). The SOEP has two main advantages. First, its longitudinal nature allows me to observe individuals' health outcomes as soon as they turn seventeen years old and match their information with household/parental variables. Second, I can observe whether individuals have attended formal childcare at some point in their life, a special characteristic of the SOEP that many administrative datasets do not have.

For the execution of the empirical strategy, I rely on an intention-to-treat design, given that I do not observe the exact age at which children attend childcare for the first time. Hence, I regress health lifestyle outcomes on the eligibility to attend formal childcare at the age of three, according to the cutoff rule. My findings yield no statistically significant effects with respect to health lifestyle outcomes later in life. Yet, one caveat of my approach is that the cutoff rule inducing children to attend childcare earlier implies that it may also induce children to be among the youngest in their elementary school cohort. To address potential violations of the exclusion restriction, I discuss the threats that this cutoff rule introduces for my identification strategy. Next, relying on recent findings from Cornelissen et al. (2016) and van den Berg and Siflinger (2016), I consider the role that family background might play in my setting. Evidence from Felfe and Lalive (2015) and Cornelissen et al. (2016) seems to suggest that there is selection with respect to the type of families that send their children to childcare at an earlier age. The descriptive evidence I have at hand also hints towards this direction. Children coming from better-off families are more likely to be enrolled in childcare earlier (Cornelissen et al., 2016). Therefore, the treatment effect for this selected sample might be negligible, as investments prior to entering childcare might already be sizeable.

Each chapter of this dissertation includes its own introduction and can be read independently. The appendices of all chapters are collected in the last section and they are arranged in the same order as the individual chapters. They contain supplementary analyses either in the form of figures or tables. The bibliography for all the three studies follows at the end of the dissertation.

CHAPTER 1

Bismarck's Health Insurance and the Mortality Decline[1]

"Rarely, if ever, in modern history has a single piece of legislation had such a profound worldwide impact as the German Sickness Insurance Law of 1883 - the cornerstone of German healthcare policy for almost one century." (Leichter, 1979)

1.1 Introduction

Improving population health is a well-accepted social objective whose impact on development and economic growth is still debated (see Acemoglu and Johnson, 2007; Bleakley, 2007; Weil, 2007; Ashraf et al., 2008; Lorentzen et al., 2008; Clark and Cummins, 2009; Jayachandran and Lleras-Muney, 2009; Hansen and Lønstrup, 2015). To estimate the effect of health on productivity and growth, most studies focus on major improvements in health and life expectancy. Such improvements take center stage during the first phase of the demographic transition at the end of the 19th century. This level shift in health, which has been coined the Epidemiological Transition, encompasses a period of marked decline in mortality particularly from infectious diseases.

The ongoing debate on the sources of this unprecedented decline in mortality that many industrializing countries experienced at that time has sprawled a range of seminal contributions. Some argue that improvements in nutrition played a crucial role (McKeown, 1979; Fogel, 2004). Others stress the role of directed public health interventions such as the roll-out of sanitation infrastructure and the diffusion of new health knowledge (Pre-

[1] This chapter is based on joint work with Stefan Bauernschuster and Erik Hornung.

ston, 1975; Szreter, 1988; Deaton, 2013).[2] So far, little is known about the role that access to health insurance played during the Epidemiological Transition.[3]

This paper investigates the role of improving health of the first ever widely implemented compulsory health insurance in the world, introduced in the German Empire in 1884. To our knowledge, we are the first to empirically evaluate the impact on mortality of the *Health Insurance Act* passed by Chancellor Otto von Bismarck as part of his social legislation program (henceforth Bismarck's Health Insurance or BHI). From December 1st, 1884, the statutory health insurance became compulsory for all industrial wage earners. Subsequently, it acted as a blueprint for Germany's current health system and served as a role model for many health systems across the world.

Using administrative data from Prussia — the largest of the German states — we present a multi-layered empirical approach that accumulates in a stepwise manner evidence for the mortality effects of Bismarck's Health Insurance and the potential mechanisms at work.[4] After looking at the development of mortality in Prussia around the introduction of BHI, we investigate Prussia's mortality decline relative to the experience of other Western European countries. Next, we exploit the fact that BHI was mandatory for blue collar workers but not for other occupations such as public servants. Newly digitized administrative data from Prussian districts allow us to compute occupation-specific mortality rates on an annual basis from 1877 to 1900. We bring the panel data to a generalized difference-in-differences model, in which we compare the mortality trend of blue collar workers (treatment group) to the mortality trend of public servants (control group) while allowing for heterogeneous reform effects over time. Finally, to address concerns regarding potential selection into occupation, we employ county and district fixed effects models that exploit regional differences in the blue-collar workers' share *before* the introduction of BHI to estimate the health insurance effects on mortality.

These different empirical approaches yield a consistent pattern suggesting negative effects of Bismarck's Health Insurance on mortality. Long-run time series data of mortality show that the beginning of the mortality decline in 19th century Prussia coincides with the in-

[2]For recent surveys see Cutler et al. (2006) and Costa (2015).

[3]Cross-country studies by Winegarden and Murray (1998) and Bowblis (2010) suggest that the expansion of health insurance across Europe is negatively associated with mortality. The exact channels through which the insurance schemes affected health outcomes remain unclear.

[4]By focusing on mortality rather than morbidity as the outcome of interest, we avoid any moral hazard and principal agent problems that might arise after the introduction of BHI.

troduction of BHI. Moreover, while mortality decreased in all Western European countries at the end of the 19th century, Prussia's mortality decline was particularly pronounced. Difference-in-differences estimates based on occupation-specific mortality rates indicate that from its introduction in 1884 to the turn of the century, BHI reduced blue-collar workers' mortality penalty by 7.8 percent. The results are robust when we allow for heterogeneous effects of urbanization, the establishment of waterworks and the roll-out of sewerage. Common pre-treatment mortality trends across occupations corroborate the validity of the identification strategy. Additionally, we find that BHI created substantial spillovers from the insured to their uninsured family members. The results are not confounded by selection into the blue-collar sector after the introduction of BHI. Using data on causes of death in a district fixed effect framework, we provide further evidence that the BHI effect is not biased by improvements in sanitation infrastructure and a resulting decline of waterborne diseases such as Typhoid fever. Rather, it turns out that a large part of the effect is driven by a reduction of mortality due to airborne infectious diseases, especially tuberculosis.

Surprisingly, the insurance was able to reduce deaths from infectious diseases in the absence of effective medication for many of the prevailing infectious diseases. These findings are in line with earlier conjectures in the historical literature arguing that the insurance contributed to the mortality decline by providing its members with access to new knowledge regarding hygiene and transmission of infectious diseases (see Koch, 1901; Ewald, 1914; Condrau, 2000). The introduction of BHI enabled families of poor workers — a group formerly unable to afford consultation from physicians — to gain access to health information and knowledge regarding hygiene and related topics. The sick funds launched information campaigns and encouraged the contracted physicians to disseminate this newly earned knowledge relating to the transmission of infectious diseases. One explicit aim of these campaigns was to bring down the number of tuberculosis incidences and it seems that these measures were indeed effective. We provide additional support for this interpretation by showing that expenditures for doctor visits are negatively related to mortality in a district fixed effects model, while sick pay and expenditures for medication and hospitalization are not.

More recent expansions of compulsory health insurance have been highly effective towards increasing access to healthcare and reducing mortality — at least for specific subgroups of the population. In this respect, our findings are in line with studies on major expansions in health insurance coverage in the U.S., such as Medicare for the elderly (see Finkelstein, 2007; Card et al., 2008; Finkelstein and McKnight, 2008; Card et al., 2009) and Medicaid for the poor (see Currie and Gruber, 1996; Goodman-Bacon, 2015). However, there are key differences between these 20th century expansions and the introduction of BHI. Current health insurance schemes work in an environment of chronic diseases and have the ability to provide healthcare by medical treatment. In contrast, BHI worked in an environment of infectious diseases without effective healthcare by medical treatment. According to Vögele (1998, p. 199-208), in the absence of proper medication, BHI might have been able to increase health by preventing families to fall into poverty due to sick pay, changing the role of doctors, increasing access to hospitals, or by allowing the state to systematically educate and control the covered population with respect to their attention to health issues. As such, BHI could work through a pecuniary channel by smoothing income and providing nutrition during times of hardship. Alternatively, in the absence of effective medication, it could work through a knowledge channel by providing access to physicians and midwifes, thus disseminating new knowledge on hygiene and the prevention of infectious diseases. To this day, econometric evidence that disentangles these channels remains largely missing.

The remainder of the paper is organized as follows. Chapter 1.2 discusses the literature on the causes of the 19th century mortality decline and provides background information on Bismarck's Health Insurance. Chapter 1.3 introduces the historical Prussian district-level data that we use in the empirical analysis. Chapter 1.4 lays out the multi-layered empirical approach, presents the results and provides a set of robustness and validity checks. Chapter 1.5 concludes.

1.2 Historical and Institutional Background

1.2.1 The Mortality Decline

Germany's mortality and fertility decline at the end of the 19th century is considered to be rather representative of the demographic transition of many European countries (Guinnane, 2011). Yet, Germany experienced higher levels of mortality and fertility decline than most other countries at the beginning of the 19th century. With respect to mortality, Leonard and Ljungberg (2010) even speak of a 'German penalty' which was particularly evident in 1870. In the beginning of the 20th century however other countries had matched the mortality levels seen in Germany. Knowledge about the precise factors that might have caused the observed mortality decline remains vague. The empirical literature has attempted to provide various explanations, ranging from nutritional improvements and public investments in sanitation infrastructure to the diffusion of knowledge as tools against the fight of disease transmission.

In 1885, life expectancy was 39.4 years for those under the age of one and 44.5 years for those between the age of 20 and 25 (Imhof, 1994[2005]). Interestingly, life expectancy in rural areas of Europe was considerably higher than in urban areas until the beginning of the 20th century (e.g., Kesztenbaum and Rosenthal, 2011). In 1877, for example, life expectancy at birth was five years less for a boy born in a Prussian city compared to a boy born in the countryside (Vögele, 1998). This 'urban penalty' was gradually removed through a range of public interventions taking place during the second half of the 19th century. Many recent studies provide evidence for major reductions in infant and adult mortality which are associated with improvements in sanitation infrastructure (Meeker, 1974; Hennock, 2000). More specifically, improvements in the water supply (Brown, 1988; Ferrie and Troesken, 2008; Beach et al., 2016), water purification (Cutler and Miller, 2005) and sewerage systems (Alsan and Goldin, 2015; Kesztenbaum and Rosenthal, 2016) strongly reduced mortality from waterborne diseases.

The seminal contributions of McKeown (1979) and Fogel (2004) conclude that improvements in living standards and nutrition were responsible for the mortality decline in 19th century Europe. Their analysis left little merit for public health interventions or medication as potential drivers of the decline. Indeed, it is widely agreed that medical treatments hardly contributed to the mortality decline before 1914 (Sieveking, 1903; Leonard and Ljungberg, 2010). Smallpox vaccination is a notable exception as it became available in 1796 and hence largely reduced infant mortality (Hennock, 1998; Ager et al., 2014). In the first decades of the 19th century, many German states introduced compulsory vaccination against smallpox. While this was not true in the case of Prussia, smallpox vaccination and re-vaccination were a widespread practice and eventually became compulsory for all children as part of the *Imperial German Vaccination Law* in 1874. Although effective medication was hardly available, scientific medical knowledge considerably increased at the end of the 19th century.

The common belief — particularly prevalent since medieval times — that diseases were transmitted through bad smells (*miasmas*) faded out in the course of the 19th century. It was instead replaced by scientific findings identifying the role of bacteria as crucial transmitters of diseases. As a result, general hygiene became a key issue in discussions about the so-called 'workers' hygiene' in Germany. Doctors and factory owners alike focused on disciplining workers in terms of sanitation with methods such as combing the hair and taking cold baths which aimed to make them more robust (Frevert, 1981; Tennstedt, 1983). Initially, it was the German elite that made efforts to improve workers' hygiene. Yet, these types of practices were gradually transmitted to the wider population as the volume of scientific evidence increased. Major breakthroughs in epidemiology and bacteriology occurred during the second half of the 19th century. These included the well-known discoveries of water as a transmitter of Cholera by John Snow and William Budd as well as numerous discoveries in bacteriology by Robert Koch, Louis Pasteur, Ignaz Semmelweis and others.[5]

Advances in bacteriology had an impact on established knowledge across all types of infections related to waterborne and airborne diseases. In fact, Mokyr (2000, p. 15) recognises germ theory to be "one of the most significant technological breakthroughs in history." However, mere identification of the root cause of infections was insufficient to cure the sick, especially when no remedies were available. All that physicians could do was to "educate patients on hygiene" (Thomasson, 2013, p. 177).

The role that knowledge diffusion played in improving health has recently gained the attention of economists. Deaton (2013), for instance, argues that upward shifts of the Preston curve are driven by the *application* of new knowledge. All around the developed world, new knowledge regarding hygiene was disseminated via healthcare centers, congresses and public information events at the end of the 19th and the beginning of the 20th century. Nordic countries were particularly progressive in diffusing knowledge through well-child visits and healthcare centers. Wüst (2012) finds that Danish infant mortality was reduced due to home visiting programs which helped towards diffusing knowledge on nutrition as well as highlighting the positive health effects of breastfeeding. Bhalotra et al. (2015) remark that a similar Swedish program managed to reduce chronic

[5] The role of hygiene, as an important tool to prevent infectious diseases in hospitals, became generally appreciated in the 1880s, twenty years after the death of Ignaz Semmelweis, the pioneer of modern antisepsis (Murken, 1983). Yet, it was not before Alexander Fleming discovered penicillin in 1928 that antibiotics became widely known as drugs that could fight bacteria.

diseases of infants by providing nutritional information, non-financial support and monitoring. According to findings by Bütikofer et al. (2015), Norwegian healthcare centers also contributed to better long-term economic outcomes by providing home visits during the first year of life. Ogasawara and Kobayashi (2015) find similar effects when evaluating a program for inter-war Tokyo. Hansen et al. (2016) investigate the introduction of tuberculosis dispensaries in Denmark in the early 20th century. In the absence of a cure, these dispensaries had, among other things, the function to disseminate knowledge related to the transmission of tuberculosis both to patients and to their families. In addition, these dispensaries had introduced information campaigns across Denmark.[6]

According to Vögele (1998), health insurance funds might have contributed to the penetration of new knowledge and health education in the German Empire. Guinnane (2003, p. 45) supports this hypothesis by noting that the insurance sickness funds played a major role in strengthening the role of physicians as advocates of hygiene. Kintner (1985) argues that contracted physicians and midwives represented a major source for disseminating information of the health effects of breastfeeding to pregnant women.[7] Further, Tennstedt (1983, p. 461) brings forward the argument that health insurance prevented the spread of illness by introducing new rules and benefits with respect to: the workers' families, the workers' hygiene and lifestyle, the workplace and the employers' duties.

Earlier work by Winegarden and Murray (1998) provides evidence showing that health insurance coverage across five European countries was associated with mortality reductions.[8] They find that a ten percentage point increase in the insured population results in a mortality reduction of 0.9 to 1.6 per 1,000 inhabitants. While issues of endogeneity in insurance take-up are likely to remain unresolved in this study, it provides an interesting yardstick to our findings. Bowblis (2010) extends this study to eleven countries to study the effect of health insurance on infant mortality. He speculates that health insurance

[6]Condran and Crimmins-Gardner (1978) argue that 'similar' information campaigns in U.S. cities at the end of the 19th century played a minor role in reducing tuberculosis-driven mortality.

[7]According to Kintner (1985) and Kintner (1987), breastfeeding was more widespread in Prussia compared to the south of Germany. In addition, while some cities in the south such as Baden or Munich saw an increase in breastfeeding, this was not the case in Berlin where breastfeeding massively declined between 1885 and 1910. Unfortunately for us, comprehensive data on breastfeeding before 1910 is missing. On another note, it has been argued that improvements in the quality and supply of cow milk were only marginal in improving infant mortality conditions (Vögele, 1998). Legal regulations regarding the quality of milk did not become an issue in the German Empire until 1901.

[8]This finding is supported by Strittmatter and Sunde (2013) who show that reductions in mortality following the introduction of public healthcare systems across Europe translate into positive effects both on income growth per capita as well as aggregate income. Their estimations are based on data from twelve European countries but exclude Germany due to lack of available data.

reduced mortality by "educating people about the benefits of clean houses, not re-using dirty bath water, washing hands and isolation of sick family members from the rest of the household" (Bowblis, 2010, p. 223). While these articles exploit across-country, over-time variation, our empirical setup exploits *within*-country, over-time, across-occupation variation. This setup allows us to flexibly control for general mortality trends within a country at the end of the 19th century. Moreover, we will take several steps to explore the potential channels of the effects.

1.2.2 Bismarck's Health Insurance

The *Compulsory Health Insurance Act* of 1883 constituted the birth of Germany's social security system. Bismarck's Health Insurance was the first of the three main branches of the German Social Insurance System, followed by the *Accident Insurance Act* (1884) and the *Disability/Old-age Pension System Act* (1891).[9] Being the first ever implemented compulsory health insurance scheme in the world, it acted not only as a blueprint for Germany's current health system but it also served as a role-model for many subsequent health insurance systems.

The decision of Chancellor Otto von Bismarck to introduce the compulsory health insurance was a reluctant reaction to mounting upheavals among the working class. The Industrial Revolution led to increasing social tension between the rising working class and the political and economic elite. The new Socialist Worker's Party of Germany gained support among the lower strata of the population and became a threat to the political stability of conservative dominance in the German Reich Parliament. Against this backdrop, the health insurance reform was a *mass bribery* for Bismarck to win over votes from the socialist party and the worker unions (Rosenberg, 1967). Furthermore, it disburdened public funds by shifting the burden of poor relief on the workers and employers. The Reichstag approved the law on May 31, 1883, against the votes of the Social Democrats who argued that this social reform would not really improve the workers' situation (Tennstedt, 1983).

[9] Fenge and Scheubel (2014) show that the introduction of the disability and old age pension system reduced fertility, while Guinnane and Streb (2011) provides evidence for moral hazard effects of the accidence insurance. Lehmann-Hasemeyer and Streb (2016) find that Bismarck's social security system as a whole crowded out private savings.

From December 1st, 1884, BHI was "compulsory for all industrial wage earners (i.e. manual laborers) in factories, ironworks, mines, ship-building yards and similar workplaces" (Act of June 15, 1883, see Leichter, 1979). Contributions were earnings-related, amounted to an average of 1.5 percent of the wage[10] and were paid jointly by employers (one-third) and employees (two-thirds). Other occupations, including public servants, farmers, domestic servants, day-laborers or self-employed were not eligible for BHI.

Benefits of BHI included free routine medical and dental care, prescribed medicine, incidental care for up to 13 weeks and treatment in hospitals for up to 26 weeks. In addition, the insurance provided maternity benefits encompassing free medical attention and a cash benefit (*Wochenhilfe*) for up to three weeks after giving birth. In the case of an insured worker's death, the insurance paid a death grant to the worker's family. Moreover, the insured were eligible to receive sick pay amounting to at least 50 percent of the average local wage for 13 weeks. Note that the national law only specified maximum contributions and minimum benefits. Thus, the individual sickness funds had "considerable discretion to set specific benefits and contribution levels" (Leichter, 1979, p. 123).

The health insurance system was administered in a decentralized manner by local sickness funds (*Krankenkassen*). Generally, we can distinguish between six types of sickness funds. Where possible, Bismarck built upon previously existing organizations such as the building trade, the miners (*Knappschaften*), the guild and various industrial sickness funds. This saved both time as well as state resources but it also made sense from a political perspective because it respected the guilds' and unions' position as insurance providers for their members. In addition to these four types of funds, two new types were established: these included the local funds (*Ortskrankenkassen*) and the parish funds (*Gemeindekrankenkassen*) whose task was to insure all eligible workers not covered by other funds. These two new funds attracted the lion's share of the newly insured workers after the 1884 reform. Indeed, evidence from Leichter (1979) suggests that, in 1905, 59 percent of all insured individuals were insured in either local or parish funds.

After issuing the Act in 1883, municipalities and other institutions had more than a year of preparatory time to set up the insurance funds. Yet, the very early period of BHI did not pass without frictions. Employers did not report their workers to the funds, workers opted to remain in pre-existing funds with lower benefits and collectors of insurance

[10]Contributions were confined to a maximum range of three to six percent of the wage.

premia often returned drunk having lost their lists.[11] Some of the workers preferred to buy insurance from voluntary funds (*Hilfskassen*) which provided lower benefits until legal adjustments were made in 1892. Voluntary funds did not require contributions by the employer, which was sometimes a good enough reason for employers to prefer to hire workers with voluntary insurance (Tennstedt, 1983, pp. 318-322). Initially, around 40 percent of the targeted workers took up insurance. In the following years, this share gradually increased – also due to more rigorous inspection.

Figure 1.1 depicts the share of the health insured in the total population over time.[12] Pre-1885 insured are either voluntarily insured or clustered in very few industries providing compulsory health insurance such as mining. The data suggest only slight increases in the insured population until 1874, i.e. the latest available pre-BHI year. Yet, after 1885, it is evident that the insured population triples. The subsequent accelerated increase in insured population is possibly attributable to several reasons such as: the increased uptake after circumventing initial frictions in the recording process of the eligible workers, an expansion in the blue-collar worker population due to ongoing industrialization and a stepwise extension of BHI towards white-collar groups (*Angestellte*).

Historical accounts propose that being covered by insurance increased the demand for health goods and services. The insured, for instance, consulted physicians far earlier and more frequently than the uninsured. Furthermore, it is argued that a large share of the newly insured would not have been able to afford to consult a physician in the absence of BHI (Huerkamp, 1985, p. 202).[13] Contracted physicians usually received a lump sum payment of 2 Marks per insured from the insurance funds, irrespective of the frequency of treatment. As a consequence, the insured increasingly made use of consultation hours. Soon, the stereotypical doctor's complaint that patients came for consultation only when it was too late had turned into complaints that patients came in

[11] Based on the occupation census of 1882 officials in Dresden were expecting 45,000 workers from 8,665 firms to be liable to compulsory insurance. By mid-1885, it turned out that only 30,000 workers were registered and 3,000 employers had yet to report their workers. Similar information is available from Leipzig (Tennstedt, 1983, p.319).

[12] Table A.3.1 in the Appendix shows the exact numbers of insured and total population of Prussia over the years.

[13] The chairman of the *Imperial Insurance Agency* Tonio Bödiker argued that less than half of the workers' families would have consulted a doctor before the introduction of the compulsory insurance (Huerkamp, 1985, pp.207-208).

for petty indispositions (Huerkamp, 1985, p. 201).[14] Moreover, BHI became a key driver of the increased utilization of hospital capacity in the 1890s (Spree, 1996).

Figure 1.1: Expansion of Health Insurance in Prussia

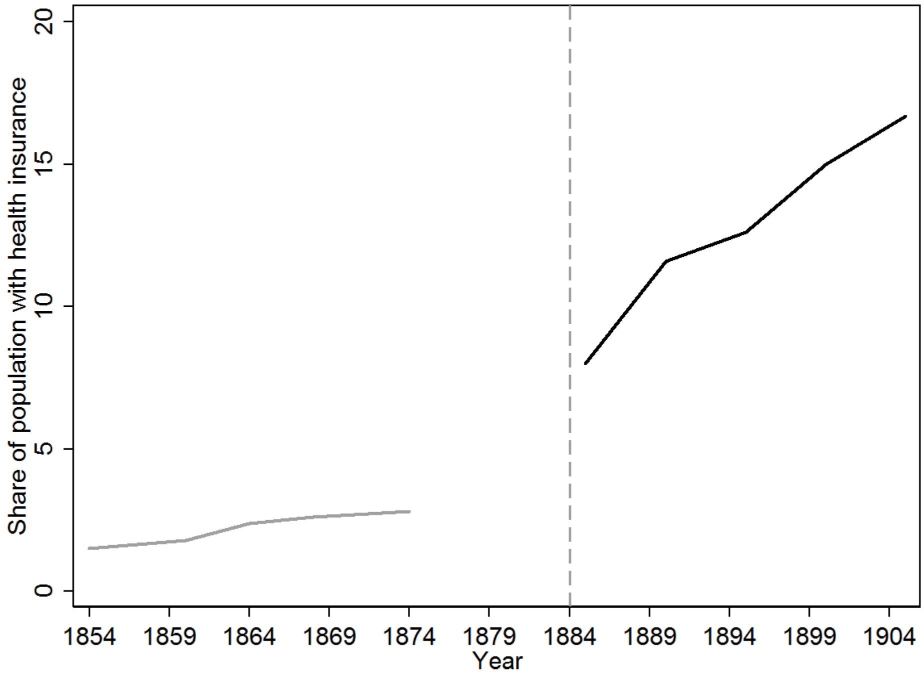
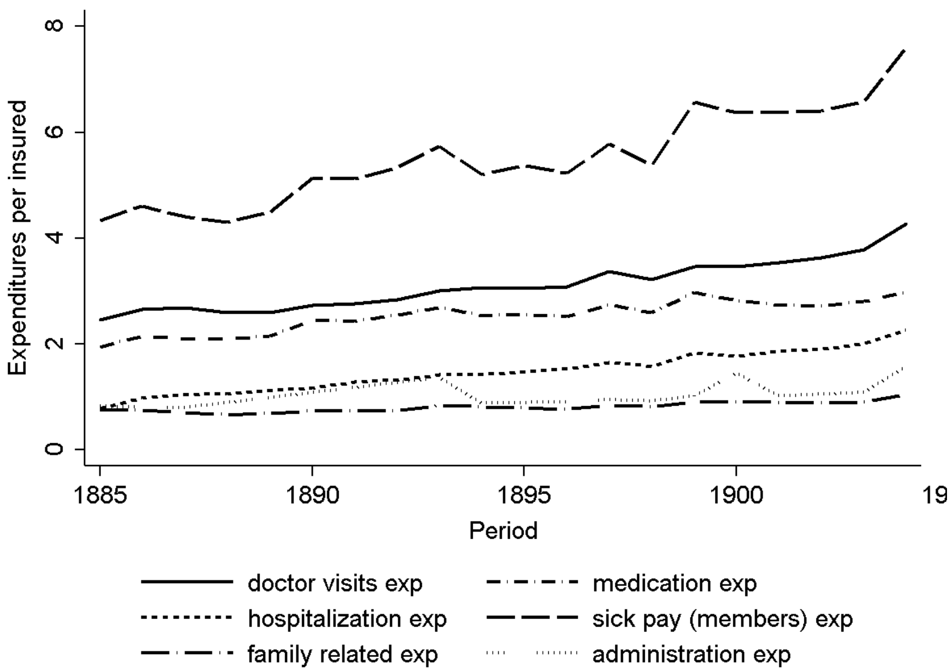


Figure 1.2: Sickness Funds' Expenditures per Insured in Marks



[14]This notion is supported by contemporary sources that suggest that only half of the consultations justified a period of sick leave (Huerkamp, 1985, pp.202).

Figure 1.2 presents the sickness funds' expenditures per insured from 1885 to 1905. While we observe a steady increase in expenditures per insured over the full period of observation, the relative importance of the different kinds of expenditures is remarkably stable. Roughly a third — and thus the largest share — of total expenditures is due to sick payments for the insured. Expenditures for doctor visits make up for another 20-25 percent of total expenditures, followed by expenditures for medication, for hospitalization and for the insured's family members. The increase in the share of insured together with the growing expenditures per insured from 1885 to 1905 suggest that any effects of BHI are likely to become stronger over time.

From a political perspective, supporters of BHI argue that the reform was costly but that it "bought social peace for Germany" (Leichter, 1979, p. 124). However, there were also more critical voices arguing that Bismarck's reform delayed the introduction of any major safety and health factory regulation (Hennock, 2007). In this paper, we would like to analyse – for the first time – the causal effect that Bismarck's reform eventually had on mortality.

1.3 Data

To quantify the effect of BHI on mortality, we draw on unique administrative data from Prussia — the largest state of the German Empire. By 1885, Prussia's territory covered roughly two thirds of the total area and population of the German Empire. The Royal Prussian Statistical Office reports the number of deceased by occupational groups. We combine these data with the Prussian population and occupation censuses. The dataset is additionally extended to include information on public sanitation infrastructure such as waterworks and sewerage. When further refining our analysis, we analyse heterogeneity in the causes of death to provide evidence on potential channels through which the reform affected mortality. The period of observation for the main analyses covers the years from 1875 to 1905.

1.3.1 Data on Mortality by Occupation

The *Preussische Statistik* (Königliches Statistisches Bureau in Berlin, 1861-1934) — a series of statistical volumes with administrative data published by the Royal Prussian Statistical Office — reports the number of deceased for all 36 Prussian districts. Starting in 1877, this information is provided annually for 28 occupational groups each of which can be further broken down by gender and by adults and children below fourteen years of age. Children and non-employed females are classified by the occupation of their father or husband, respectively.

The 28 occupation categories range from blue-collar sectors such as metals, textiles, chemicals and foods to public service sectors such as education, healthcare and public administration, but also include farming and fishing or the trade sector. This categorization is particularly helpful for our analysis since it allows us to distinguish between mortality of occupational groups that are eligible for compulsory health insurance and mortality of those that are not.

The period under analysis is a period of rapid industrialization in Prussia.[15] Accordingly, the occupational groups may have experienced differences in the growth of the working population leading to differences in the growth of the population at risk (see Table A.3.2 for details). To take this into account, we use data from occupation censuses that were conducted in the years 1882, 1895 and 1907 under the supervision of the *Imperial Statistical Office* (Kaiserlich Statistisches Amt, 1884-1942). We linearly extrapolate and interpolate the data between the censuses to obtain estimates of the respective size of each occupation group in the years with missing data. The main occupational categories in the occupation censuses nicely match those reported in the mortality statistics of the Royal Prussian Statistical Office. Also similar to the mortality statistics, the occupation censuses report the occupation-specific number of dependent children and non-working wives. This allows us to not only compute occupation-specific total death rates but also to distinguish between males, females and children.[16]

1.3.2 Data on Hospitalizations, Causes of Death, Health Insurance Contributors and Aggregate Mortality

Additionally, we use rich data on hospitalization rates and causes of death from the Königliches Statistisches Bureau in Berlin (1861-1934), which were digitized for the purpose of this project. In particular, district level information on the number of treated

[15]Note that this period is not a period of substantial warfare. Deaths related to the Franco Prussian war of 1871 are unlikely to substantially change the mortality pattern of the period 1877-1900.

[16]We tested the robustness of our results to replacing this denominator by the occupation-specific working population. The results are qualitatively similar.

patients and cases per year is available starting in 1880. Moreover, starting in 1875, the number of deceased for 30 distinct causes of death is reported annually until 1902. We can for example distinguish between deaths from waterborne infectious diseases such as Typhus, Typhoid fever, or Diarrhea and deaths from airborne infectious diseases such as smallpox, scarlet fever, measles, diphtheria, pertussis, scrofula, tuberculosis, tracheitis, or pneumonia.[17]

Moreover, we draw on data recording the actual take-up of health insurance. The Kaiserlich Statistisches Amt (1884-1942) provides annual records on the number of insurance contributors by district. Although these data are available for the entire German Empire, we confine our analysis to Prussia since our detailed mortality data are only available for the Prussian territory.

Finally, Galloway (2007) provides aggregate data on mortality for all 441 Prussian counties for each year from 1875 to 1905. Apart from informing on total mortality, the data distinguish between male and female mortality and also between mortality of legitimate and illegitimate infants. Infant mortality is defined as the number of deaths of children below the age of one divided by the number of live-births in thousands during the same year. This applies for legitimate and illegitimate deaths and births respectively.

1.3.3 Data on Urbanization and Public Sanitation Infrastructure

We address concerns regarding changes specific to the urban environment coinciding with the introduction of the health insurance by controlling for urbanization rates and sanitation infrastructure. Population censuses were conducted every five years in Prussia. We use all population censuses available for our period of observation to compute urbanization rates (Galloway, 2007). The literature discusses two main drivers of changes in urban mortality occurring at the end of the nineteenth century which are related to the provision

[17]Concerns regarding the quality of causes of death data in this period have been raised in the literature (Kintner, 1999; Lee et al., 2007). It is likely that improved knowledge of diseases allowed registrars to better identify the accurate cause of death over the course of our period of observation. If regions with a higher share of insured were also regions were physicians or registrars were better able to identify the cause of death, we would expect ‘unknown cause of death’ to show a stronger decline in these regions. Using the category ‘unknown cause of death’ as an outcome variable of a district fixed effects model, we do not find any systematic relation between this variable and the blue-collar workers’ share in 1882. This finding indicates that there were no systematic improvements in the ability to identify the correct cause of death related to the introduction of BHI that could drive our findings on causes of death.

of public sanitation infrastructure — waterworks and sewerage. We digitized data on waterworks from Grahn (1898-1902) and on sewerage from Salomon (1906-1907) reporting the year in which Prussian cities started to provide public water supply and sewerage to population. Assuming that the entire city population benefited from the introduction of the sanitation infrastructure, we calculate the share of the total urban population in a county and a district with access to public waterworks and sewerage on an annual basis.

1.4 Empirical Evidence

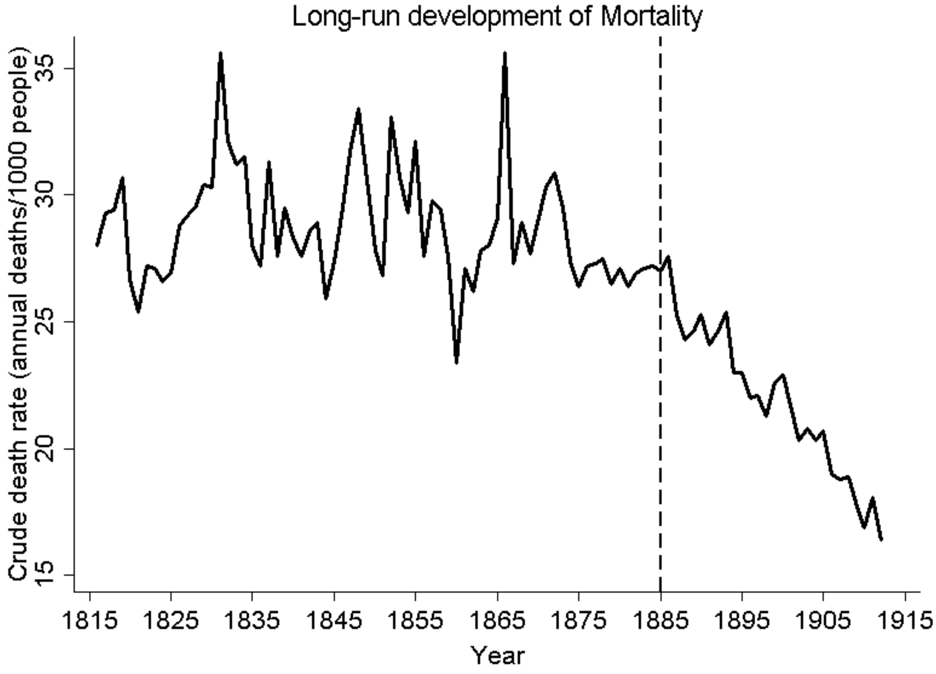
This section empirically analyses the relationship between the introduction of BHI and mortality. To pin down the relationship of interest, we take a multi-layered approach which aims at addressing various concerns by bringing together evidence from different datasets and from varying degrees of aggregation. The section gradually builds up specifications starting from country-level time-series data, to an intention-to-treat difference-in-differences design with occupation-specific mortality rates and finally county and district level designs which address selection issues and allow us to be more specific regarding the channel through which BHI affected mortality rates. Each subsection is structured to first lay down the econometric specification, then to present the results and finally to discuss advantages, concerns and drawbacks specific to each approach.

1.4.1 Time-series and Cross-country Statistics

We start our empirical analysis by inspecting the long-run development of mortality in Prussia from the early 19th to the early 20th century. Figure 1.3 plots the crude death rate, defined as the number of deaths per 1,000 inhabitants of Prussia over the period 1815-1913. Mortality was rather volatile until the early 1870s when the fluctuations notably ceased. However, it was not before the mid-1880s that we observe a distinct break in the long-run mortality trend. From 1885 to 1913, the crude death rate in Prussia declined from about 27 to about 17 deaths per 1,000 inhabitants, corresponding to a substantial drop of almost 40 percent. Thus, we observe a remarkable coincidence of the introduction of BHI in December 1884 with the timing of the mortality decline.

But was the Prussian mortality decline from 1885 to 1915 also special from an international perspective? To tackle this question, we compute mortality rates of selected European countries using data on deaths and population size from a range of national sources that

Figure 1.3: Long-run Development of Mortality in Prussia



are collected and made available by the team of the Human Mortality Database.[18] In Figure 1.4, we plot the crude death rates for Prussia and various European countries against years from 1875 to 1913. To smooth the trends, we apply country-specific local regressions using a tricube weighting function (Cleveland, 1979) and a bandwidth of 0.15. Prussia's mortality rate was comparatively high at the time when Bismarck's Health Insurance was introduced. From 1884 to 1913, mortality rates declined across all countries. Yet, there is hardly any country in which the mortality decline was as pronounced as in Prussia. To highlight this fact, we plot the difference in the mortality rate of Prussia and every other country by year, while normalizing to zero the respective mortality difference in 1884. Again, we apply local regressions using a tricube weighting function and a bandwidth of 0.15 to smooth the time trends. Figure 1.5 shows that the mortality decline for Prussia was indeed considerably stronger than the mortality decline of all other countries during this period. Only the Netherlands experienced a comparably strong decline. Although remarkable, we should not interpret these simple time-series and cross-country statistics as causal evidence for an effect of BHI on mortality. There might be structural changes coinciding with BHI that did not affect mortality in other countries but are responsible for Prussia's comparatively strong mortality decline at the end of the 19th century. There-

[18]For details, please visit http://www.mortality.org/.

fore, in the remainder of this chapter, we will put together additional pieces of evidence to plausibly separate the effect of the health insurance from that of other determinants of mortality.

Figure 1.4: Mortality Decline across Western European Countries

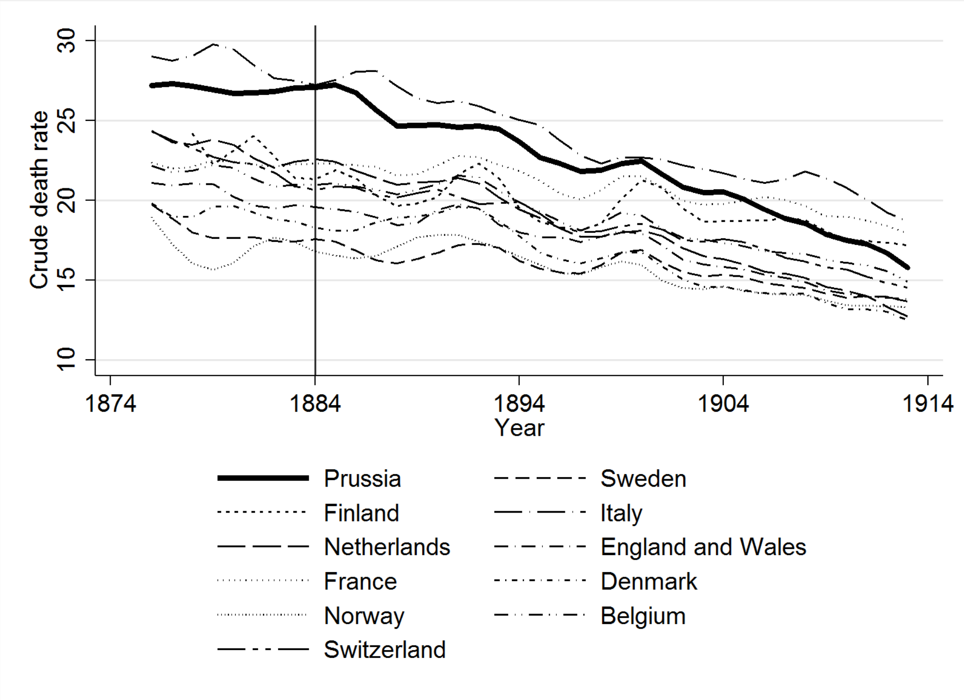
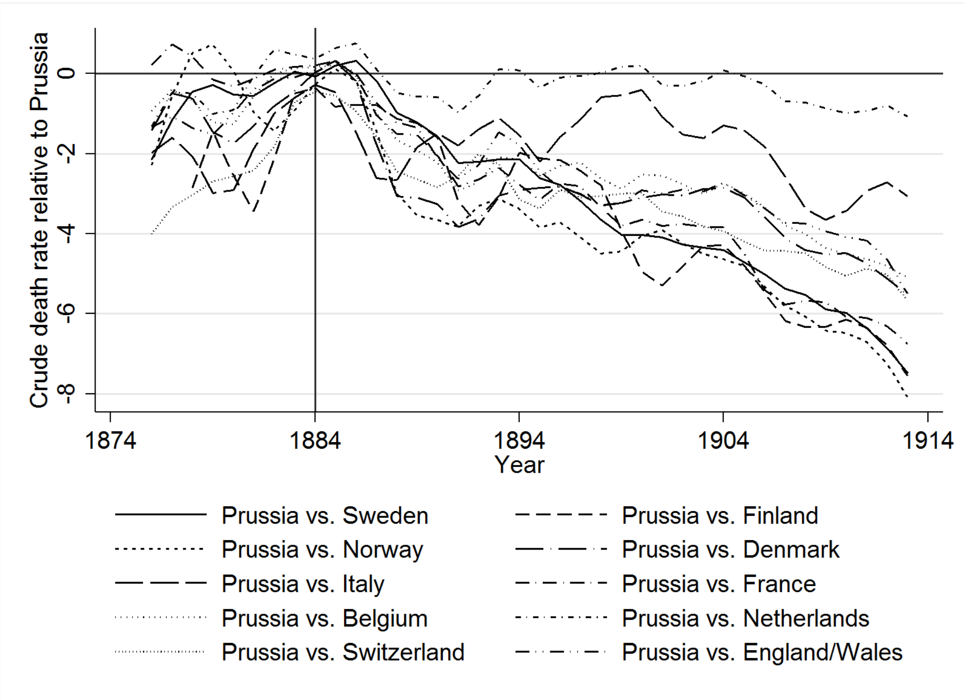


Figure 1.5: Mortality Decline relative to Prussia



1.4.2 Difference-in-Differences: Eligibility by Occupation

Econometric specification

To investigate the role of Bismarck's Health Insurance on mortality, we proceed by exploiting the fact that BHI — introduced in December 1884 — was mandatory for blue collar workers but not for other occupations. This constitutes a natural setting for a reduced-form difference-in-differences model in which we compare the mortality trend of blue collar workers (treatment group) to the mortality trend of public servants (control group).

Two characteristics qualify public servants as our preferred control group. First, similarly to blue collar workers, public servants are likely to live in urban areas and thus experience the same structural changes to their living environment. Second, public servants did not become eligible for compulsory health insurance before 1914. According to the *Imperial Law on the Legal Relationship with Public Servants* of 1873, public employees were eligible for continuation of salary payment during illness and a pension in case of disability or old age. They did not however receive benefits such as free doctor visits and medication. Finally, most importantly, this situation for public servants did not change between 1873 and 1914.

Exploiting this fact, we estimate a difference-in-differences model that can be expressed by the following equation:

$$Death_{iot} = \alpha_{io} + \theta_{it} + \sum_{t=1877-1880}^{1897-1900} \beta_t BlueCollar_{io} + X'_{it} BlueCollar_{io}\gamma + \varepsilon_{iot} \qquad (1.1)$$

$Death_{iot}$ is the average death rate of people with occupation $o \in (BlueCollar, PublicServant)$, measured in district i in period $t \in (1877-1880, 1881-1884, 1885-1888, 1889-1892, 1893-1896, 1897-1900)$.[19] α_{io} are occupation by district fixed effects accounting for any time-constant occupation-specific mortality differences between districts. θ_{it} are district by period fixed effects that flexibly allow mortality trends to differ across districts. Thus, these fixed effects pick up a range of shocks affecting the

[19] The use of four year periods is owed to the fact that the occupation-specific mortality was published from 1877 — eight years before BHI. We are thus able to create two pre-treatment periods and four post treatment periods. Results are robust to the choice of other period lengths and the use of annual data (see Section 1.4.2).

district-level health environment, i.e. both occupational groups equally. These could be overall improvements in nutrition due to variation in harvests and food prices, or differences in temperature especially affecting infant mortality. $BlueCollar_{io}$ is a dichotomous variable that is unity for blue collar workers and zero for public servants. β_t measures the reduced-form effect of BHI if there are no time-varying unobservables that affect blue-collar workers' mortality differently from public servants' mortality. ε_{iot} is a mean zero error term. Standard errors are clustered at the district level to allow for serial autocorrelation within districts.

By letting β vary over time, we generalize the standard difference-in-differences model to allow for heterogeneous intention-to-treat effects over time. This makes particular sense in our setting where we expect the mortality effects of BHI to expand gradually. At the same time, this specification allows us to perform a placebo treatment test. In particular, if we use the period from 1881-1884 as the omitted category and find β_t to be zero in the pre-treatment years, this suggests that blue collar workers and public servants indeed followed the same mortality trend before BHI was introduced. Thus, this placebo treatment test would corroborate the validity of our identifying assumption, namely that the mortality of blue collar workers and public servants follow the same time trend in absence of the treatment.

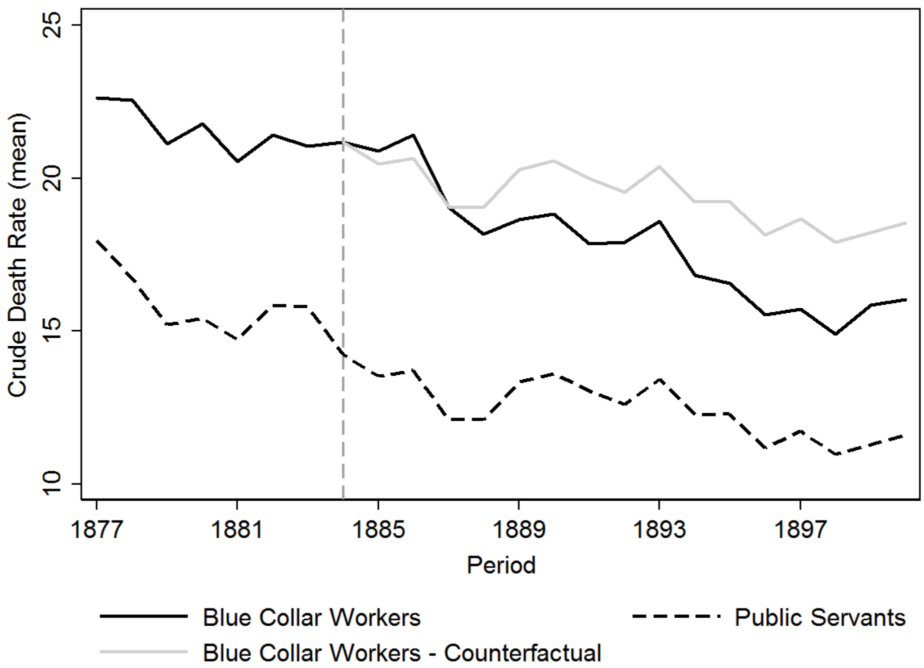
To further validate our empirical approach, we introduce — in an extended specification — an interaction of the blue-collar worker dummy $BlueCollar_{io}$ with a vector of time-varying district-level control variables X'_{it}. Public health interventions, such as the construction of waterworks and sewerage in cities, are among the most frequently cited explanations for decreased mortality in 19th century Europe and also in the U.S. (see Ferrie and Troesken, 2008; Alsan and Goldin, 2015; Beach et al., 2016; Kesztenbaum and Rosenthal, 2016). Accordingly, X'_{it} includes the district's urbanization rate, the share of a district's urban population with access to public waterworks and the share of a district's urban population with access to public sewerage. It is important to mention that time-varying district-level characteristics are already captured by the θ_{it} in our basic specification as long as they affect both occupational groups equally. However, if the effects vary by occupation, they might still confound the estimates. Therefore, in the extended specification, we explicitly allow measures of public health infrastructure to have different effects for blue collar workers and public servants. Furthermore, by allowing urbanization rates to differentially affect occupational groups, we account for the fact that city quarters with occupational clustering could have been differentially affected by changes in population density due to city growth at the intensive margin.

Main results

A first graphical depiction of the difference-in-differences setup is provided in Figure 1.6. Here, we plot the crude death rate of blue collar workers (treatment group, black solid line) and the crude death rate of public servants (control group, black dotted line) against years. The vertical line marks the introduction of BHI in 1884. In addition, the grey solid line depicts the counterfactual mortality trend of blue collar workers, i.e. the mortality trend they would have followed without BHI under the assumption that the mortality trend of public servant gives us an image of what would have happened to blue-collar workers' mortality without the treatment. Throughout the entire period of observation, the crude death rate of blue collar workers lies above the crude death rate of public servants. In the years before BHI, both groups follow approximately the same mortality trend. If anything, public servants' mortality is even declining faster than blue-collar workers' mortality, which would rather downward bias the difference-in-differences estimate. Only after the introduction of BHI, the mortality of blue collar workers is falling more steeply than the mortality of public servants. This is evident from the large deviation of blue-collar workers' actual mortality trend from the counterfactual trend. We interpret this graphic pattern as suggestive of a negative treatment effect of BHI on the mortality of blue collar workers.

In a next step, we bring the data to a regression framework and estimate the generalized difference-in-differences model of Equation (1.1). Column 1 of Table 1.1 reports the results from a basic specification, where we regress the crude death rate on the interactions of the blue-collar worker dummy and period fixed effects while controlling for district by occupation fixed effects and district by period fixed effects. The period immediately preceding Bismarck's reform, i.e. the period 1881-1884, constitutes the omitted category. We find that blue collar workers and public servants indeed followed the same mortality trend in the years preceding BHI. This result provides supportive evidence for the common trend assumption of the difference-in-differences framework and thus corroborates the validity of the empirical approach. After 1884, the crude death rate of blue collar workers first increases as compared to the death rate of public servants. This short-lived

Figure 1.6: Crude Death Rates: Blue Collar vs. Public Servants



deterioration of blue-collar workers' health might be related to the adverse health environment of blue collar workers and the initial frictions of the BHI introduction. For all subsequent periods, we observe highly significant negative effects that gradually increase in size. By the end of the 19th century, BHI had reduced the mortality penalty of blue collar workers by 1.654 deaths per 1,000 individuals, or by 7.8 percent (-1.654/21.184).

To account for occupation-specific urbanization effects, i.e. the crowding of factory workers into city quarters due to rapid city growth, we now include the interaction of the urbanization rate with a blue-collar worker dummy as a covariate. The results from column 2 show that this slightly reduces the point coefficients. Yet, the effects stay negative, statistically significant and economically meaningful. The same is true if we include occupation-specific interactions of access to waterworks (column 3) or access to sewerage (column 4) to make sure that the results are not confounded by potentially occupation-specific effects of the roll-out of sanitation infrastructure. Across all specifications, the results point to a negative and significant effect of BHI on blue-collar workers' mortality, which increases over time.

Figures A.1.1 and A.1.2 in the Appendix provide further graphical support for the argument that the roll-out of waterworks and sewerage does not confound the health insurance effect. They show the number of waterworks and sewerages established in Prussian cities

Table 1.1: Flexible DiD: Main Results

Dep. var.: Deaths per occ. Pop.	Base (1)	Urban (2)	Waterworks (3)	Sewerage (4)
BlueCollar x 1877	-0.1905	-0.2767	-0.3233	-0.2004
	(0.2495)	(0.2784)	(0.2720)	(0.2515)
BlueCollar x 1885	1.1325***	1.2109***	1.2217***	1.1590***
	(0.1616)	(0.1475)	(0.1603)	(0.1595)
BlueCollar x 1889	-0.7339***	-0.5881**	-0.4385*	-0.6635**
	(0.2573)	(0.2310)	(0.2536)	(0.2658)
BlueCollar x 1893	-1.3014***	-1.0734***	-0.8695***	-1.1875***
	(0.2560)	(0.2724)	(0.2886)	(0.3065)
BlueCollar x 1897	-1.6540***	-1.2496**	-1.1373***	-1.4888***
	(0.2842)	(0.4798)	(0.3366)	(0.4069)
Urbanization x Occupation	No	Yes	No	No
Waterworks x Occupation	No	No	Yes	No
Sewerage x Occupation	No	No	No	Yes
District x Occupation FE	Yes	Yes	Yes	Yes
District x Time FE	Yes	Yes	Yes	Yes
Observations	432	432	432	432
R-squared	0.94	0.94	0.94	0.94

Notes: Table reports flexible DiD estimates. All variables are averaged over four year periods from 1877-1900. Dependent variable measures crude deaths rates using deaths by occupation of the household head per alive occupational population (including dependents) in thousands. The omitted period is 1881-84. Standard errors clustered at the district level in parentheses. * 10%, **5%, *** 1% confidence level

per year as well as their cumulative distribution functions. Both waterworks and sewerage coverage in Prussian cities clearly increases in the second half of the 19th century. However, we do not observe any conspicuous jump in the roll-out around the introduction of Bismarck's Health Insurance in 1884 which could explain the absolute and relative mortality decline of blue collar workers in the aftermath of the reform.

Effect heterogeneity: men, women and children

So far, we have looked at the average reduced-form effect of BHI on the crude death rate. In order to obtain more information on the underlying components of this effect, we disaggregate the occupation-specific death rates to obtain separate death rates for men, women and children. Children and non-employed females are classified by the occupation of their father or husband respectively. Table 1.2 presents the results of this exercise. We find a large part of the mortality decline to be driven by male blue collar workers (column 2). At the same time, we observe a substantial decrease of child mortality (column 4),

Table 1.2: Flexible DiD: Heterogeneity

Dep. var.: Deaths per occ. Pop.	Base (1)	Adult Male (2)	Adult Female (3)	Kids (4)
BlueCollar x 1877	-0.1905	-0.2996	-0.7517***	0.0292
	(0.2495)	(0.2822)	(0.1640)	(0.1689)
BlueCollar x 1885	1.1325***	0.9058***	0.1703	0.6533***
	(0.1616)	(0.1570)	(0.1074)	(0.1167)
BlueCollar x 1889	-0.7339***	-0.6812***	-0.2988*	-0.2516
	(0.2573)	(0.2044)	(0.1538)	(0.1648)
BlueCollar x 1893	-1.3014***	-0.9128***	-0.3604**	-0.7350***
	(0.2560)	(0.2205)	(0.1644)	(0.1692)
BlueCollar x 1897	-1.6540***	-1.3008***	-0.5787***	-0.9260***
	(0.2842)	(0.2668)	(0.1962)	(0.2094)
District x Occupation FE	Yes	Yes	Yes	Yes
District x Time FE	Yes	Yes	Yes	Yes
Observations	432	432	432	432
R-squared	0.94	0.91	0.89	0.93

Notes: Table reports flexible DiD estimates. All variables are averaged over four year periods from 1877-1900. Dependent variable measures crude deaths rates using deaths by occupation of the household head per alive occupational population (including dependents) in thousands. *Adult Male and Adult Female* use occupational working population by gender as denominator. The omitted period is 1881-84. Standard errors clustered at the district level in parentheses. * 10%, **5%, *** 1% confidence level

while the effect on females is somewhat smaller but also significantly different from zero (column 3).[20] For all three groups, the effects gradually increase over time.[21]

Looking at the economic literature on early human capital development and the fetal origin hypothesis (Deaton, 2007; Douglas and Currie, 2011), we should not be surprised to see that children respond to changes in the health environment although they are not themselves insured. In particular, we expect children to benefit from BHI via intra-family spillovers. These intra-family spillovers might for example be induced by sick payments. If sick payments stabilize family income, they might facilitate continuous calorie intake, enhance nutritional prospects of the family and thus also be beneficial for infant health (see Subramanian and Deaton, 1996; Case and Paxson, 2008). Alternatively, access to healthcare for the insured might have an effect on the health of the entire family. In the absence of effective medication, spillovers potentially occur through the prevention

[20]Note that instead of using the full occupation-specific population, regressions for adult males and females use the occupation-specific working population. Especially for the male population, this is arguably very close to capturing only the insured population in the treatment group. Results are similar when using the full occupation-specific population as denominators.

[21]In further unreported regressions, we checked that our findings are robust to excluding Berlin, Prussia's biggest city and a district itself, from the sample.

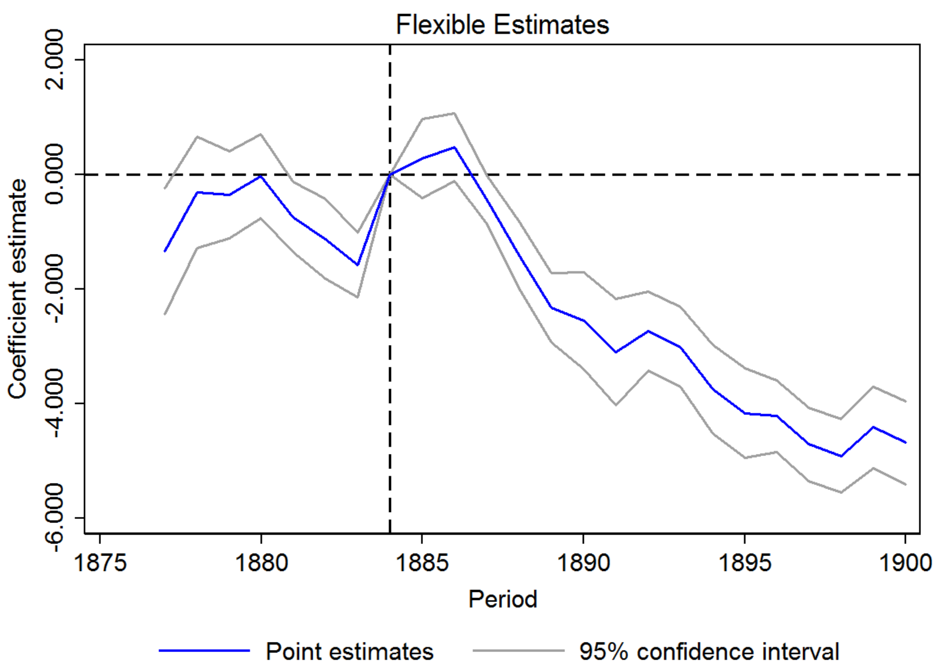
of contagion of infectious diseases. As indicated above, contracted physicians provided individuals — in particular female household heads — with knowledge on personal hygiene matters, which should reduce the incidence of infectious diseases at the household level and consequently children's mortality.

Robustness checks

To receive a more nuanced view of the effects of BHI, we estimate a difference-in-differences model in which we substitute the four-year blocks by single years. In this case, we estimate a basic version of the model without the health infrastructure and urbanization covariates. However, to test the robustness of our findings, we control for potentially different pre-treatment trends between blue collar workers and public servants.

Figure 1.7 plots the annual difference-in-differences estimates using year 1884 as the omitted category. In line with previous findings, we observe a slight increase in mortality right after the reform. However, we can now see that the mortality reduction kicks in immediately after 1886. In the years that follow, mortality of blue collar workers gradually declines when compared to mortality of public servants. It turns out that allowing for diverging pre-treatment trends makes the intention-to-treat estimates even larger than those presented in Table 1.1.

Figure 1.7: Annual DiD Estimates



Achieving an accentuated drop in mortality is easier if the pre-existing level is high. Blue-collar workers' mortality was indeed higher than public servants' mortality before the introduction of BHI. Therefore, the observed relative reduction of blue-collar workers' mortality might just constitute regression to the mean. Yet, we argue that a similar convergence would not have been achieved in the absence of access to health insurance for blue collar workers. To provide further evidence for this interpretation, we now use crude death rates measured on a logarithmic scale as an alternative dependent variable. The results confirm the established pattern, confirming that our estimates do not merely capture regression to the mean effects (see Figure A.2.2 in the Appendix).

Threats to identification

Bismarck's disability insurance and old age pension system

Bismarck's disability insurance and old age pension system, the third pillar of the welfare system, was introduced in 1891, i.e. seven years after BHI. Using 1891 as the baseline year in the annual difference-in-differences model, we investigate whether the introduction of the third pillar constitutes a considerable trend break. The estimates show that relative mortality starts to decline long before 1891 and continues to do so (see Figure A.2.1 in the Appendix). Indeed, there is no particular pattern in the data suggesting that the year 1891 changed blue-collar workers' relative mortality in a meaningful way. This finding takes away concerns that the disability insurance and the old age pension system confound our effects from Bismarck's Health Insurance.

Factory regulation and working conditions

One might be concerned that meaningful improvements in working conditions for blue collar workers just coincide with the introduction of BHI. If this was true, these improvements might lead to a stronger mortality decline for blue collar workers than for public servants and thus confound the BHI effects. However, note that the period under analysis is a period of ongoing, rapid industrialization. This is why we expect deterioration rather than improvement in industrial working conditions. Moreover, the *Trade, Commerce and Industry Regulation Act* (*Gewerbeordnung*) of 1878, which barred children under twelve from working in factories, mines, foundries and stamping mills, marked "the end of the development of factory legislation in Germany for the next thirteen years" (Hennock, 2007, p.83).

Bismarck strongly opposed any further attempts aimed at improving working conditions since he considered new factory regulations to be detrimental to economic development. Hennock (2007) argues that Bismarck's Health Insurance might have even delayed any major safety and health regulations in factories. Indeed, the 1880s saw only very few improvements in workplace regulation. The few federal regulations were rather minor and restricted to very specific industries. In particular, these improvements were related to the following regulations: the use of white phosphorus in the manufacture of matches (1884), the manufacture of lead paints and lead acetate (1886) and the manufacture of hand-rolled cigars (1888). Almost immediately after Bismarck resigned in 1890, regulations were passed that reduced maximum working hours for women. Yet, in the end, men could not benefit from any working-hour regulations before 1919 (e.g. Hennock, 2007, pp.125-128). Thus, it seems unlikely that these minor improvements could drive the BHI effects. Still, we will later present further empirical evidence against improved working conditions as a confounder. Specifically, we use data on causes of death to show that the negative mortality effect is not driven by a reduction of accidents in the workplace.

A related concern might be that blue-collar workers' income grew more rapidly than public servants' income and that this confounds the estimates. To be precise, a relative increase of blue-collar workers' income would confound the estimates as long as the reason for the relative increase is not BHI. If income is affected by BHI on the other hand, income is a channel for the BHI effect on mortality. As a consequence, we should not control for income in our regressions since this would give rise to selection bias arising from bad control. To analyse how relevant this concern is, we collected data on income dynamics in our period of observation. Geisenberger (1972, p.183) draws on data from the Royal Prussian Statistical Office to show that industrial wages increased by around 14 percent from 1885 to 1895. We use data from the Royal Prussian Statistical Office to analyse teachers' wages and also find a 14 percent increase from 1886 to 1896.[22] If we accept teachers' wages as a proxy for public servants' wages, these statistics suggest that public servants' wages increased by roughly the same rate as blue-collar workers' wages. Consequently, differential income dynamics for blue collar workers and public servants cannot explain the relatively strong mortality decline of the former.

[22] To make the data consistent over time, we compute teachers' wages net of seniority payments (*Alterszulage*) and the value of accommodation provided by the employer.

Spillovers and selection

As indicated above, positive spillovers within the family of the insured are very likely — especially if reductions of infectious diseases drive the mortality decline. Yet, since families do not live in isolation, spillovers might not just occur within the family but also affect other individuals outside the family. This might be particularly relevant in our setting. On the one hand, we would like our control group to be as similar as possible to our treatment group, i.e. they should for example live in the same area. On the other hand, this means that the control group might benefit from the improved disease environment. Spillovers from blue collar workers to public servants due to reduced risk of contagion would imply that our estimates constitute a lower bound of the BHI effects on mortality.

Moreover, the data at hand do not allow a clear distinction between insured and uninsured people, neither in the treatment nor in the control group. We cannot exclude the possibility that public servants buy voluntary health insurance. At the same time, not all blue collar workers might be insured due to frictions in particular during the early years of BHI. Consequently, there might be people in the treatment group who are uninsured and people in the control group who are insured.[23] Therefore, we should interpret our approach as an intention-to-treat design that identifies reduced-form effects of BHI eligibility on mortality. On the positive side, this approach rules out any issues arising from the endogenous nature of actual insurance take up.

Whereas selection into insurance take-up should not be a problem in our setup, selection into blue-collar occupations might bias the estimates. The key problem is that the data do not allow us to fix assignment to treatment to a date before the introduction of BHI. We only observe the full occupational population at each point in time. This gives rise to concerns about systematic selection into treatment, i.e. people selecting into blue-collar occupations whose health characteristics are systematically different from those in the treatment group before the introduction of BHI. It might be that young and healthy people from rural areas migrate to cities to pick up an industrial occupation. As a result, the average age structure of the treatment group might decrease which would lead to lower mortality. On the other hand, new workers from rural areas might particularly suffer from the dismal living conditions of urban working-class quarters and fall sick leading to higher

[23] Moreover, within both occupational groups we cannot distinguish between dependent workers and the self-employed who would not be captured by the mandatory nature of BHI.

mortality. Thus, the direction of the bias arising from selection into the industrial sector is a priori unclear.

1.4.3 Fixed Effects: Pre-Reform Differences at the County Level

In this section, we aim at resolving the discussed spillover and selection issues using an alternative empirical specification. This specification allows us to estimate the effect of BHI on the full population, capturing all potential spillovers. At the same time, the alternative specification enables us to exclude selection into treatment eligibility by holding fixed the treatment group at a point in time before the introduction of BHI.

Econometric specification

To this end, we use a county fixed effects model and compare the mortality trend of counties with a high share of blue collar workers at the time BHI was introduced to the mortality trends of counties with a lower share.[24] This model can be described by the following equation:

$$Death_{it} = \alpha_i + \theta_t + \sum_{t=1875-1879}^{1900-1904} \beta_t BlueCollar_{i,1882} + X'_{it}\gamma + \varepsilon_{it} \tag{1.2}$$

Y_{it} is the average death rate of county i in year-block $t \in (1875-1879, 1880-1884, 1885-1889, 1890-1894, 1895-1899, 1900-1904)$. α_i are county fixed effects capturing unobserved time-invariant heterogeneity between counties, and θ_t are year-block fixed effects that flexibly account for general time trends. $BlueCollar_{i,1882}$ is the share of blue collar workers in county i in year 1882. We hold this variable constant at the 1882 level and interact it with period dummies. Thus, β_t captures any year-block specific associations between the share of blue collar workers in 1882 and the outcome variable. By holding the share of blue collar workers constant at the 1882 level, we avoid problems related to potentially systematic selection into the industrial sector after the introduction of BHI. X'_{it} is a vector of time-varying county-level covariates, including the urbanization rate, the share of population with access to sewerage and waterworks and the initial crude mortality rate for the period 1880-1884 interacted with period dummies. Controlling for interactions of initial mortality and period dummies accounts for the fact that counties

[24]The county is the administrative unit below the district. Prussia consists of 441 counties with an average area of less than 800 square kilometers.

differ in their levels of mortality prior to the introduction of BHI. Therefore, we account for the fact that counties might follow different mortality trajectories over time. ε_{it} is a mean-zero error component. Standard errors are clustered at the county level.

In order for β_t to identify reduced-form intention-to-treat effects of BHI on mortality, we rely on the assumption that there are no time-varying unobserved determinants of mortality that have different effects in counties with a high share of blue collar workers in 1882 than in counties with a lower share. In other words, these counties follow the same mortality trend in absence of the treatment. Again, we provide evidence in support of this assumption by performing a placebo treatment test, i.e. we analyse whether β_t is indeed zero in the pre-treatment years. Additionally, we employ a placebo treatment group to test the validity of this alternative approach.

Main results

In column 1 of Table 1.3, we document that the negative and significant Bismarck effect on mortality can be replicated using a basic version of the county fixed effects model, as described in Equation (1.2). Similarly, the reduced-form effect of Bismarck's Health Insurance increases over time. In the average county (with a blue-collar share of 7.4 percent in 1882), BHI reduced mortality by roughly 0.9 deaths per 1,000 inhabitants, measured at the turn of the century. The insignificant placebo treatment effect in the pre-treatment period corroborates the validity of this alternative empirical approach. In column 2, we add the urbanization rate, the share of population with access to waterworks and the share of population with access to sewerage to control for changes in urbanization patterns and the roll-out of public health infrastructure. Moreover, we include interactions of the 1880-1884 mortality rate with period dummies to allow for mortality trends that might differ between counties depending on their initial mortality levels. Adding these controls does not change the findings qualitatively.[25] Yet, the reduced-form effects become slightly smaller as waterworks and sewerage pick up part of the reduction in mortality over time.

In columns 3 to 7 of Table 1.3, we disaggregate the outcome variable to analyse the effects of BHI on males, females, and infants separately.[26] We find statistically significant negative effects for males and females after 1885. For infants, the post-BHI coefficients

[25]Findings are qualitatively similar if we interact the urbanization and public infrastructure variables with time dummies to allow for differential effects over time (results available upon request).

[26]Infant mortality is defined as deaths of children during their first year of life per 1,000 births.

Table 1.3: County Fixed Effects using 1882 Blue-collar Workers' Share

	Base (1)	Controls (2)	Male (3)	Female (4)	Infants (5)	LegInf (6)	IllegInf (7)	Placebo (8)
Treatment x 1875	1.891	1.326	-6.225***	0.303	-19.852*	-22.991**	12.609	-1.961
	(1.623)	(1.615)	(1.668)	(1.754)	(10.882)	(11.504)	(56.893)	(4.792)
Treatment x 1885	-7.768***	-6.381***	-6.101***	-5.638***	9.342	6.694	39.778	0.406
	(1.693)	(1.656)	(1.949)	(1.582)	(11.245)	(11.772)	(46.841)	(2.912)
Treatment x 1890	-4.754**	-2.072	-4.540*	-4.658**	-11.106	-23.222	62.624	-1.100
	(2.041)	(2.023)	(2.368)	(2.087)	(14.944)	(15.522)	(52.273)	(3.619)
Treatment x 1895	-8.631***	-5.333**	-2.858	-6.164***	-20.445	-42.638**	106.611	-1.113
	(2.080)	(2.101)	(2.556)	(2.180)	(19.035)	(19.820)	(72.781)	(4.067)
Treatment x 1900	-12.330***	-8.572***	-12.522***	-10.346***	-25.992	-55.566**	142.091**	3.313
	(2.399)	(2.330)	(2.727)	(2.516)	(23.081)	(23.656)	(69.946)	(4.798)
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2645	2645	2645	2645	2645	2645	2645	2645
Counties	441	441	441	441	441	441	441	441
Periods	6	6	6	6	6	6	6	6
R-squared	0.75	0.77	0.76	0.75	0.28	0.30	0.06	0.77

Notes: Table reports county-level fixed effects estimates. All variables are averaged over five year periods from 1875-1905. Dependent variable measures crude death rates using total deaths per alive population in thousands. Treatment variable is the blue-collar workers' share observed in 1882, interacted with time-period dummies. Column *Controls* adds the urbanization rate, waterworks per capita, sewerage per capita, and initial mortality of the period 1880-1884 interacted with time-period dummies. Dependent Variable in Column *Male* is male mortality. Dependent Variable in Column *Female* is female mortality. Dependent Variable in Column *Infants* is infant mortality (< 1 year) per 1,000 births. Dependent Variable in Column *LegInf* is infant mortality (< 1 year) born in wedlock per 1,000 births in wedlock; Dependent Variable in Column *IllegInf* is infant mortality (< 1 year) born out of wedlock per 1,000 births out of wedlock. Dependent Variable in Column *Placebo* is total mortality; treatment variable is public servants share observed in 1882, interacted with time-period dummies. Standard errors, clustered at the county level, in parenthesis. * 10%, **5%, *** 1% confidence level are negative but insignificant. Yet, once we distinguish between the effects on legitimate infants (column 6) and illegitimate infants (column 7), we find significantly negative effects for the former. Thus, mortality reductions due to BHI are only found for infants that grow up in households with both parents present. The fact that there are no BHI effects for illegitimate infants might be explained by the lack of intra-family spillovers for this group of infants: if a single mother is not a blue collar worker, the absence of a father leaves the family without access to insurance benefits. Finally, note that in those cases where the pre-treatment interaction is significant, the sign is always negative. Consequently, if we accounted for these deviating trends in the pre-treatment period, the treatment effects would become even larger.

Column 8 presents an additional placebo test. Here, we test the hypothesis that differences in counties' pre-BHI share of public servants (as a placebo treatment group) are able to create a similar pattern of results as the share of blue collar workers in a county. We reject this hypothesis as we find only very small and insignificant coefficients for the relationship between initial public servants share and changes in mortality. Consequently, we can rule out that our treatment indicator picks up mortality trends common to other occupational groups.[27]

Robustness checks

In a next step, we move to the district level since this level of regional aggregation allows us to use new data on the share of individuals registered with health insurance funds. First, we slightly modify Equation (1.2) to reproduce the county level results for the district level (columns 1-5). Then, we use the share of individuals registered with a health insurance fund instead of the share of blue collar workers as the key variable of interest. Similarly to the previous specification, we hold the share of health insured constant in 1885, right after the introduction of BHI and interact it with year dummies. This ensures that the results are not confounded by selective insurance take-up in the years following the introduction of BHI. If the results for the share of health insured and the share of blue collar workers are similar, this provides evidence that our findings are indeed related to insurance and not driven by unrelated, unobserved improvements in blue-collar work. In fact, columns 6-10 of Table 1.4 show that the results are very similar to the ones we obtain in columns 1-5.

[27] Note that running a horse race regression between the share of blue collar workers and the share of public servants yields similar results (available upon request).

Table 1.4: District Fixed Effects using Blue Collar Workers and Insured Population

Dep. var.: Deaths per Pop.	Initial Blue Collar Workers (1882)					Initial Insured (1885)				
	Base	Controls	Male	Female	LegInf	Base	Controls	Male	Female	LegInf
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment x 1875	3.369	3.147	-3.970	1.806	-1.339	5.042	4.871	-3.711	3.907	7.588
	(3.355)	(4.651)	(4.494)	(5.580)	(50.455)	(4.299)	(5.408)	(5.054)	(6.734)	(56.659)
Treatment x 1885	-11.191**	-10.892**	-11.603*	-11.795**	-19.468	-15.005**	-14.620**	-14.849*	-15.185**	-41.337
	(4.603)	(5.211)	(5.748)	(5.184)	(40.938)	(6.114)	(6.950)	(7.848)	(7.068)	(53.348)
Treatment x 1890	-7.075	-7.014	-11.626	-12.110	-84.091	-11.274	-11.188	-18.050	-17.289	-132.667
	(7.896)	(7.996)	(10.588)	(8.960)	(75.537)	(10.647)	(10.515)	(13.860)	(11.887)	(93.314)
Treatment x 1895	-11.641	-11.753	-11.654	-15.248*	-151.433	-16.735*	-17.011*	-18.793	-21.503**	-219.387*
	(6.981)	(7.041)	(9.778)	(7.825)	(95.252)	(9.159)	(9.040)	(12.626)	(10.062)	(113.869)
Treatment x 1900	-15.049**	-15.001**	-19.470**	-19.116***	-170.702*	-21.831**	-21.980***	-28.648**	-27.851***	-235.432**
	(7.275)	(6.296)	(8.760)	(7.006)	(86.635)	(9.001)	(7.519)	(10.893)	(8.413)	(104.441)
Controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	216	216	216	216	216	216	216	216	216	216
Districts	36	36	36	36	36	36	36	36	36	36
Periods	6	6	6	6	6	6	6	6	6	6
R-squared	0.88	0.89	0.86	0.86	0.52	0.88	0.89	0.87	0.87	0.53

Notes: Table reports district-level fixed effects estimates. All variables are averaged over five year periods from 1875-1905. Dependent variable in Columns 1, 2, 6 and 7 is crude death rates measured as total deaths per alive population in thousands. Dependent variable in Columns 3 and 8 is male mortality. Dependent variable in Columns 4 and 9 is male mortality. Dependent variable in Columns 5 and 10 is infant mortality (< 1 year) born in wedlock per 1,000 births in wedlock. Treatment variable in columns 1-5 is blue-collar workers' share observed in 1882, interacted with time-period dummies. Treatment variable in columns 6-10 is share of the population covered by health insurance observed in 1885, interacted with time-period dummies. Controls include the urbanization rate, waterworks per capita, sewerage per capita and initial mortality of the period 1880-1884 interacted with time-period dummies. Standard errors, clustered at the district level, in parenthesis. * 10%, **5%, *** 1% confidence level

Moreover, the district level enables us to use data on hospitalizations. Since BHI provided free treatment in hospitals, we expect an increase in hospitalizations. Using the number of people seeking hospitalization as an outcome variable in a district fixed effects model along the lines of Equation (1.2), we find that the introduction of BHI is indeed associated with an increase in the number of individuals being admitted to hospitals. This finding is in line with Spree (1996) who argues that Bismarck's Health Insurance increased the utilized capacity of hospitals. Detailed results are available from the authors upon request.

1.4.4 Exploiting Data on Causes of Death and Sick Funds' Expenditures

In this section, we use panel data on causes of death to provide further evidence against contemporaneous reforms and public health improvements biasing the results. Moreover, the causes of death data allow us to better investigate the channels via which BHI reduced mortality rates. Since the causes of death data are not recorded by occupation, we cannot use it in the model of Equation (1.1). However, we can employ these in a regional fixed effects model along the lines of Equation (1.2).[28] Finally, we will present regression results using data on sick funds' expenditures that further support the interpretation of our findings.

Further evidence against confounding factors

In 1885, Otto von Bismarck introduced the second pillar of the German welfare system, namely the accident insurance. This could raise suspicions that the health insurance effect is confounded by the accident insurance reform or by resulting improvements in working conditions. To test the relevance of this concern, we use deaths by accident as the outcome variable of our district fixed effects model. As the results depicted in column 1 of Table 1.5 reveal, we do not find any negative association between the number of blue collar workers in 1882 and changes in mortality due to an accident. The coefficients are significant but positive and hence they go towards the opposite direction. This implies that the overall mortality effect is not driven by a decline of deaths by accidents. Thus, this exercise provides evidence against the accident insurance confounding the effect of Bismarck's Health Insurance on mortality. Instead, the results are in line with Hennock

[28] Note that, due to changes in the original reporting of the causes of death after 1902, the last period only contains three years of data from the period 1900-1902.

(2007) who argues that BHI delayed any major safety and health regulations in factories and with Guinnane and Streb (2012) who show that the introduction of the accident insurance resulted in an increase of work-related accidents due to moral hazard.

Despite controlling for urbanization, waterworks and sewerage, there might still be concerns that nonlinearities in the improvements in water supply might confound the estimates. To address this concern, we take a thorough look at deaths by waterborne diseases that were strongly reduced by the introduction of better water supply (Ferrie and Troesken, 2008). If our effects were driven by waterborne diseases, this would leave room for the BHI estimates being confounded by improvements in water supply. Column 2 of Table 1.5 shows that there is no association between the share of blue collar workers in 1882 and deaths by waterborne diseases. Indeed, the coefficients are far from any conventional significance levels and even positive in all of the four post-treatment periods. This finding is encouraging as it provides additional evidence which shows that changes in water supply do not confound the effect of Bismarck's Health Insurance on mortality.

Evidence for the relevance of the information channel

As indicated in Section 1.2, debates on the main determinants of the health transition at the turn of the 19th century are ongoing. Seminal contributions by McKeown (1979) and Szreter (1988) have divided the literature into proponents of economic and social changes leading to improvements in the nutrition on the one side and proponents of improvements due to public health interventions such as the roll-out of waterworks and sewerage on the other side. However, what the literature generally agrees on is that efficient medication was widely unavailable at that time. But even if doctors could not really cure people, what they could do is provide them with new information on hygiene to prevent the spread of infectious diseases (Thomasson, 2013). Thus, Bismarck's Health Insurance might have enhanced access to doctors, which in turn reduced mortality via an information channel.

As laid out above, we do not find any effects of BHI on deaths from waterborne infectious diseases. This does not come as a complete surprise. Even if individuals knew from contracted physicians about the risk of contaminated water, all they could do is boil drinking water. Apart from that, they could hardly avoid getting into contact with contaminated water unless proper sewerage systems and waterworks were installed. Thus, we might argue that any reduction in deaths from infectious waterborne diseases should

Table 1.5: District Fixed Effects: Causes of Death

Dep. var.: Deaths per Pop.	Accident	Infectious	Infectious	Infectious	Infectious	Maternal
	(1)	(2)	(3)	(4)	(5)	(6)
		Waterborne	Airborne	Lung	TB+Scrofula	
Treatment x 1875	0.153	3.229*	-3.603*	-2.775*	-0.039	0.004
	(0.164)	(1.781)	(1.872)	(1.487)	(0.644)	(0.087)
Treatment x 1885	-0.079	2.374	-8.716***	-1.633	-0.761	-0.203***
	(0.150)	(1.850)	(2.731)	(1.787)	(0.667)	(0.067)
Treatment x 1890	0.275**	1.890	-4.669	-3.996	-3.588***	-0.198**
	(0.123)	(3.065)	(4.091)	(2.802)	(1.036)	(0.095)
Treatment x 1895	0.294**	0.496	-4.887	-7.245***	-5.911***	-0.185*
	(0.127)	(4.140)	(3.676)	(2.604)	(1.363)	(0.095)
Treatment x 1900	0.425***	0.664	-7.793**	-10.763***	-8.138***	-0.035
	(0.155)	(4.006)	(2.971)	(2.260)	(1.799)	(0.124)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	216	216	216	216	216	216
Districts	36	36	36	36	36	36
Periods	6	6	6	6	6	6
R-squared	0.58	0.28	0.81	0.68	0.88	0.92

Notes: Table reports district-level fixed effects estimates. All variables are averaged over five year periods from 1875-1905. Dependent variable measures death rates by cause of deaths using total deaths per alive population in thousands. Treatment variable is blue-collar workers' share observed in 1882, interacted with time-period dummies. *Accident* is death from Accident; *Waterborne* deaths from Typhus, Typhoid fever and three types of diarrheal diseases; *Airborne* deaths from Smallpox, Scarlet Fever, Measles, Diphtheria, Pertussis, Scrofula, Tuberculosis, Tracheitis, Pneumonia and other lung diseases; *Maternal* is death in childbed. Controls include the urbanization rate, waterworks per capita, sewerage per capita, and initial mortality of the period 1880-1884 interacted with time-period dummies. Standard errors, clustered at the district level, in parenthesis. * 10%, **5%, *** 1% confidence level be investment driven. In contrast, individual hygiene measures might be more successful in reducing the risk of infection with airborne diseases.

Indeed, while BHI had no effects on deaths related to waterborne diseases, we do find negative and statistically significant effects on mortality related to airborne disease (column 3). Out of all airborne infectious diseases, tuberculosis was the most prominent at that time. For the group of 20 to 70 year olds, tuberculosis was responsible for about 30% of the deaths. A major breakthrough in fighting the disease was achieved in 1882, when Robert Koch identified the bacterium causing tuberculosis.[29] Since tuberculosis strongly affected the working age population, the sickness funds were particularly interested in reducing the incidence of this disease. Yet, as the cure was not developed until 1946, the focus was set to preventing infections. The hygienic situation of workers' housing became the center of attention. A characteristic excerpt from Tennstedt (1983, p.458) mentions that *research by Preysing and Schutz found tuberculosis germs underneath 21.2% of 66 toddler's fingernails, which they absorbed by crawling on the floors of worker dwellings contaminated by sputum.* Robert Koch (1901, p. 575) himself argued that only preventive action could reduce tuberculosis mortality, including the diffusion of knowledge about its contagiousness to increasingly larger circles.

A major role for detecting deficits and educating workers in hygiene is attributed to sickness inspectors (*Krankenkontrolleure/Krankenbesucher*). Insurance funds employed inspectors whose task was to pay unexpected home visits to monitor the curfew and medication intake of patients.[30] Moreover, sickness fund doctors tried to particularly influence the women of the household, who were typically in charge of care and food (Frevert, 1981).[31] Our results suggest that these measures were successful. As it can be seen from columns 4 and 5 of Table 1.5, the share of blue collar workers in 1882 is associated with a decline in the number of deaths by lung diseases and especially tuberculosis (and scrofula). These effects are meaningful in size, statistically significant and become gradually stronger over time.[32] On a side note, we also find some evidence for a reduction in maternal death in childbed (column 6).

[29]Initially, tuberculosis was often assumed to be hereditary since usually the entire family suffered from its symptoms.

[30]In 1896, the municipality fund of Leipzig conducted 79,332 visits by voluntary inspectors and 149,899 visits by professional inspectors (Tennstedt, 1983, p. 451).

[31]According to Tennstedt (1983, p. 458) funds considered to deploy female inspectors to give advice to *their uneducated sisters* on *how to ventilate and clean the apartment, curtains and other dust catchers.*

[32]As technology to detect the actual cause of death was limited, the deaths classified as unspecific lung diseases could in fact have been tuberculosis and vice versa.

As a final step in our exploration of the underlying channel of the effect, we now exploit data on expenditures of health insurance funds. The data can be distinguished between several types of expenditures, including: expenditures for doctor visits, medication, hospitalization, sick pay and maternity benefits. If indeed knowledge diffusion by physicians played a crucial role in reducing mortality, we would expect mortality to decrease as doctor visits increase. On the other hand, if sick pay was crucial — on the grounds that it allows for stable family income and thus avoids temporary dry spills in nutritional intake — we would expect mortality to decrease as cash benefits increase. To get an intuition of the relevance of these alternative channels, we regress district level mortality rates on the different types of expenditures of the insurance funds. District fixed effects account for time-invariant heterogeneity between districts, while time fixed effects flexibly capture mortality trends common to all districts.

Column 1 of Table 1.6 provides estimation results that lend empirical support to the knowledge diffusion channel. An increase in expenditures for doctor visits by one standard deviation decreases the mortality rate by 0.094 standard deviations. An increase in medication expenditures, however, does not affect the mortality rate (column 2). The same is true for hospitalizations (column 3) as well as sick pay to the insured and sick pay to the dependents (columns 4 and 5). Further, the negative effect of maternity benefits does not reach conventional significance levels (column 6). In columns 7 and 8, we find a positive relationship between the mortality rate and death benefits as well as expenditures for administration. This finding reminds us that we should not over-interpret these results since the estimates might suffer from endogeneity problems. Especially in the case of death benefits, causality is likely to run from mortality to death benefits and not the other way round.[33] In column 9, all types of health expenditures are simultaneously included in the model. The results of this extended model confirm the negative effects of expenditures for doctor visits on mortality, while there are no associations between medication expenditures, hospitalization, or sick pay and mortality. The negative effect of maternal benefits on mortality now also turns significant. This leaves some room for nutrition as another channel via which BHI might have affected mortality. Paid maternity leave for three weeks could smooth a family's income after birth and give the mother the

[33] An alternative interpretation is that the death of an insured person is indicative of deterioration of the health prospects of the rest of the family by reducing access to healthcare.

Table 1.6: Mortality and Health Expenditures

Dep. var.: Total Mortality Rate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Doctor visits(std)	-0.094**								-0.104**
	(0.040)								(0.042)
Medication(std)		-0.026							0.066
		(0.082)							(0.103)
Hospitalization(std)			-0.026						-0.004
			(0.056)						(0.074)
Sick pay members(std)				0.079					-0.013
				(0.070)					(0.073)
Sick pay dependents(std)					0.010				-0.000
					(0.029)				(0.049)
Maternity benefits(std)						-0.027			-0.056**
						(0.020)			(0.027)
Death benefits(std)							0.082**		0.129***
							(0.037)		(0.044)
Administration(std)								0.015***	0.008**
								(0.004)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	144	144	144	144	144	144	144	144	144
Districts	36	36	36	36	36	36	36	36	36
Periods	4	4	4	4	4	4	4	4	4
R-squared	0.92	0.91	0.91	0.92	0.91	0.92	0.92	0.91	0.93

Notes: Table reports OLS panel estimates. Dependent variable measures crude death rates using total deaths per alive population in thousands averaged over five year periods 1885-1905. Explanatory variable is sickness fund expenditure per insured measured at the beginning of each 5 year period (1885, 1890, 1895, 1900). All variables are standardized with mean zero and unit standard deviation. Controls include the urbanization rate, waterworks per capita, sewerage per capita and initial mortality of the period 1880-1884 interacted with time-period dummies. Standard errors, clustered at the district level, in parenthesis. * 10%, **5%, *** 1% confidence level possibility to stay home and start breastfeeding, which both can have positive effects in particular on newborns' health.[34]

In sum, this exercise provides further evidence that the negative mortality effect of BHI likely worked mainly through providing a new group of people with access to physicians. New knowledge on hygiene provided by physicians and sickness funds' inspectors was thus more easily diffused to a population living under poor hygienic condition. This in turn resulted in the prevention of infections of airborne diseases such as tuberculosis.

1.5 Conclusion

Which role did health insurance schemes play during the time of demographic transition? This question is interesting in many respects. First, understanding the role of public institutions for demographic change and economic growth is crucial for the design of effective public policies. Second, from a demographic perspective, there has been a lot of work on determinants of the demographic transition. Yet, the role of the first health insurance schemes has been largely neglected so far. Third, from a historical perspective, it is interesting to understand the effects of Bismarck's Health Insurance as the first compulsory health insurance scheme in the world.

We use newly digitized Prussian administrative panel data to analyse the effects on mortality of the first compulsory health insurance in the world as introduced by Otto von Bismarck, Chancellor of the German Empire, in December 1884. We start with evidence from time-series data on long-run mortality in Prussia, move on to an international comparison, and finally use difference-in-differences type frameworks that exploit the compulsory nature of the health insurance scheme for blue collar workers. The different empirical approaches yield a consistent pattern suggesting that BHI played a crucial role for Prussia's sharp mortality decline at the end of the 19th century.

In an intention-to-treat design, we find a mortality reduction of 1.654 deaths per 1,000 blue collar workers by the end of the century. In other words, blue-collar workers' mortality decreased by 7.8 percent due to the introduction of Bismarck's Health Insurance. Findings are robust to alternative specifications and placebo tests. Additional evidence suggests

[34] An alternative interpretation might of course again be reverse causality. Maternity benefits per insured are higher in districts with higher fertility. The increased number of newborns in a district increases the denominator of the mortality rate and thereby decreases the outcome of interest.

that a large part of the reduction in mortality is driven by a decline in deaths related to airborne infectious diseases which are subsequently related to wider access to physicians after the reform. This evidence is supportive of our hypothesis that the health insurance provided families at the lower end of the income distribution with access to physicians and new knowledge on hygiene matters.

If we draw parallels between the magnitude of the findings following Bismarck's Health Insurance introduction and findings from the modern U.S. health insurance literature, we observe that in the former case, the results are considerably larger. In other words, the reduced-form effects under Bismarck's reform are much larger when compared to evidence from extending eligibility for Medicare or Medicaid programs. Analytically, Bismarck's Health Insurance — by the end of the 19th century — had reduced the mortality of blue collar workers by 7.8 percent compared to baseline mortality. This is particularly striking if one considers that Bismarck's health reform was introduced during a time when effective medication was not available. In the U.S. setting, Card et al. (2008) find that extending Medicare eligibility for the 65+ group led to a mortality reduction between 2 and 4 percent one year later. In another study, Currie and Gruber (1996) find that increasing Medicaid eligibility to low-income children is associated with a 3.4 percent mortality reduction. As a final point, however, it is important to highlight that in the absence of clean information on insurance take-up, one should be cautious when comparing reduced-form intention-to-treat effects across papers.

CHAPTER 2

Gender Differences and Stereotypes in Financial Literacy: Off to an Early Start

2.1 Introduction

Financial literacy is a reliable predictor of individual wealth, savings, stock market participation and retirement planning (Lusardi and Mitchell, 2008; van Rooij et al., 2011). Yet, financial literacy is low on average, particularly among women. The gender gap in financial literacy is large and persists throughout the life cycle, but the underlying mechanisms are poorly understood. Explanations of gender gaps in other domains, such as differences in risk attitudes, self-confidence, or division of labour, can only partially account for the gap in financial literacy (Lusardi and Mitchell, 2014).

Most studies of the gender gap in financial literacy focus on adults. An exception is Lührmann et al. (2015b) who document that the gender gap in financial literacy already exists at younger ages – among 13 to 15 year old teenagers. In this paper, we report results from a field study in the same age group that elicited domain-specific measures of gender stereotypes jointly with measures of financial literacy.

Our interest in gender stereotypes is motivated by recent findings showing that stereotypes can explain gender gaps in various domains (Bordalo et al., 2016; Coffman, 2014; Lavy and Sand, 2015). In the present context, stereotypes represent beliefs about the levels of and the future returns to, the financial knowledge of women and men. Bordalo et al. (2016) present a social cognition model of stereotypical thinking that implies overreaction to information that confirms stereotypes. Stereotypical beliefs may thus lead to under-investment in financial knowledge among girls.[1]

Our data show gender differences in the relationship between the strength of gender stereotypes and the level of financial literacy among teenagers. This relationship is robust to controlling for several other factors that may explain the gender gap, such as numeracy, risk preferences (Niederle and Vesterlund, 2010) and self-confidence (Bucher-Koenen et al., 2014). Beliefs are biased towards higher competency of males: teenagers of both genders believe that boys have higher interest and ability regarding financial matters; that the returns to financial knowledge are higher for males; and that males are more likely to deal with financial matters at work. Further, there is no gender difference in financial knowledge among those teenagers who do not share male-favouring stereotypical views. The more strongly teenagers agree with such stereotypes, the wider the gender gap.

2.2 Data and Survey Instruments

We study the association between gender stereotypes and teenagers' financial knowledge in a sample of 418 high-school students recruited from 30 classes across 13 schools in three German cities.[2] Participating schools pertain to the two lower tracks of the German high school system. Students in those tracks typically continue with vocational training after graduation and come from lower socio-economic status backgrounds.

We assess financial knowledge through standard financial literacy questions on discounting, interest compounding, the time value of money, risk diversification and the definition of stocks (Lusardi and Mitchell, 2014; Lusardi et al., 2010; van Rooij et al., 2011). They are similar to those used in the PISA module on financial knowledge. We ask seven additional questions to cover concepts like return, liquidity and risk of different assets, running versus one-off costs, budgeting skills and cash versus payment in installments. We construct an index of financial knowledge as the number of correct answers to the twelve questions.

[1] Jappelli and Padula (2013) and Lusardi et al. (2016) present investment models of financial human capital.

[2] Classes were randomly drawn from a list of classes interested in a financial education course offered by a non-profit organisation. 97% of students provided participation consent signed by their parents. Data were collected in 2013. For more details, see Lührmann et al. (2015a) who analyze an experiment on time preference measurement conducted as part of the same study.

Gender stereotypes are measured in five sub-domains (Table 2.1): financial interest (or motivation), the ability to deal with financial matters, the relevance of financial knowledge at work and at home and future returns to financial literacy. Questions are answered on a five-point Likert scale. We construct an index of belief bias towards males by summing up the responses and re-scaling to the unit interval.[3]

All questionnaires were filled out in the presence of the research team during regular school hours. We asked students for their gender, age and social background: household composition, the language they speak at home and the number of books at home. The latter is a standard question in PISA, capturing important family inputs into a student's education (Hanushek and Woessmann, 2011). The survey also included math grades, self-reported risk attitude (see Dohmen et al. (2011) for empirical validation), self-confidence (using the scale proposed in Robins R. W. and Trzesniewski (2001)) and four of Raven's progressive matrices to measure cognitive skills, based on K.A Heller (1998).

2.3 Results

Female teenagers have lower financial knowledge than their male counterparts (the difference amounts to 0.3 of a standard deviation, $p = 0.0028$, Mann-Whitney-Wilcoxon (MWW) test[4]); this result is similar to Lührmann et al. (2015b). Gender differences in financial literacy may be related to gender-specific risk attitudes, numeracy and self-confidence. In columns 1 and 2 of Table 2.2, we present summary statistics for these variables and for socio-economic characteristics. Adolescent females neither have lower levels of numeracy than males (as evidenced by their last math grade)[5], nor different risk attitudes, self-confidence, or cognitive ability (column 3 in Table 2.2). Male and female teenagers in our sample are also similar in terms of their socio-demographic characteristics, such as household size, number of siblings, migrant background and number of books in the household.

[3] Psychologists often rely on implicit association tests to determine biases in beliefs and stereotypes. As no such test existed specifically for the financial domain, we developed our 5-item scale, leaning on the "Beliefs about Women Scale" by Belk and Snell (1986).

[4] Throughout, we test for gender differences using the MWW test for ordinal and continuous variables and the χ^2 test for dummy variables.

[5] We define a dummy for low numeracy which is 1 if a students' grade is below the class average (since math exams are not standardised in Germany) and 0 otherwise.

Table 2.1: Survey Instrument to Measure Stereotypes

		Females' attitudes	Males' attitudes	$H_0 : x_f = x_m$ (p-value)
Interest	*"Men are usually more interested in finances than women."*	2.85	3.15	0.0081***
Ability	*"Men are usually better in dealing with finances than women."*	2.64	3.09	0.0000***
Work	*"Men are more likely to be concerned with finance in their job than women."*	2.73	2.96	0.0523*
Home	*"Men are more likely to be concerned with the family finances than women."*	3.00	3.17	0.1132
Expected return	*"For a successful future it is more important for men to be good at dealing with finance than for women."*	2.72	3.07	0.0013**
Overall index	aggregating all answers and rescaling to $S \in (0, 1)$	0.44	0.52	0.0002***

Note: Stereotypes are measured on a 1-5 Likert scale, 1 = "not at all" and 5= "absolutely true". Higher values indicate bias towards males. p-values refer to gender differences in financial stereotypes, using MWW tests.

Table 2.2: Summary Statistics and Estimation Results

		Summary Statistics		Text	Estimation Results	
		Females	Males	$H_0 : x_f = x_m$ (p-value)	Females' Financial Knowledge	Males' Financial Knowledge
Age (months)		171.3 [9.991]	170.6[9.509]	0.562	0.012 [0.014]	0.015 [0.014]
>26 books at home	D	0.105 [0.307]	0.159 [0.366]	0.114	0.918 [0.575]	0.395 [0.321]
Household size (log)		1.057 [0.574]	1.049 [0.511]	0.605	-0.017 [0.257]	-0.184 [0.264]
Siblings	D	0.808 [0.395]	0.780 [0.415]	0.493	-0.099 [0.460]	0.019 [0.225]
Migrant Background	D	0.448 [0.499]	0.524 [0.500]	0.123	0.120 [0.294]	0.319 [0.246]
Grade 8	D	0.477 [0.501]	0.516 [0.501]	0.427	-0.181 [0.309]	1.162 [0.307]***
Low math score	D	0.436 [0.497]	0.435 [0.497]	0.982	-0.601 [0.225]**	-1.018 [0.245]***
High cognition	D	0.209 [0.408]	0.236 [0.425]	0.524	-0.046 [0.340]	0.331 [0.328]
High risk	D	0.203 [0.404]	0.260 [0.440]	0.180	0.674 [0.302]**	-0.162 [0.328]
Low confidence	D	0.192 [0.395]	0.146 [0.354]	0.217	-0.024 [0.331]	-0.277 [0.453]
City Fixed Effects					Yes	Yes
Observations		172	246	418	172	246
R-squared					0.091	0.233

Note: The dependent variable is the average no. of questions answered correctly. D indicates dummy variables. Robust standard errors clustered at the class level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1

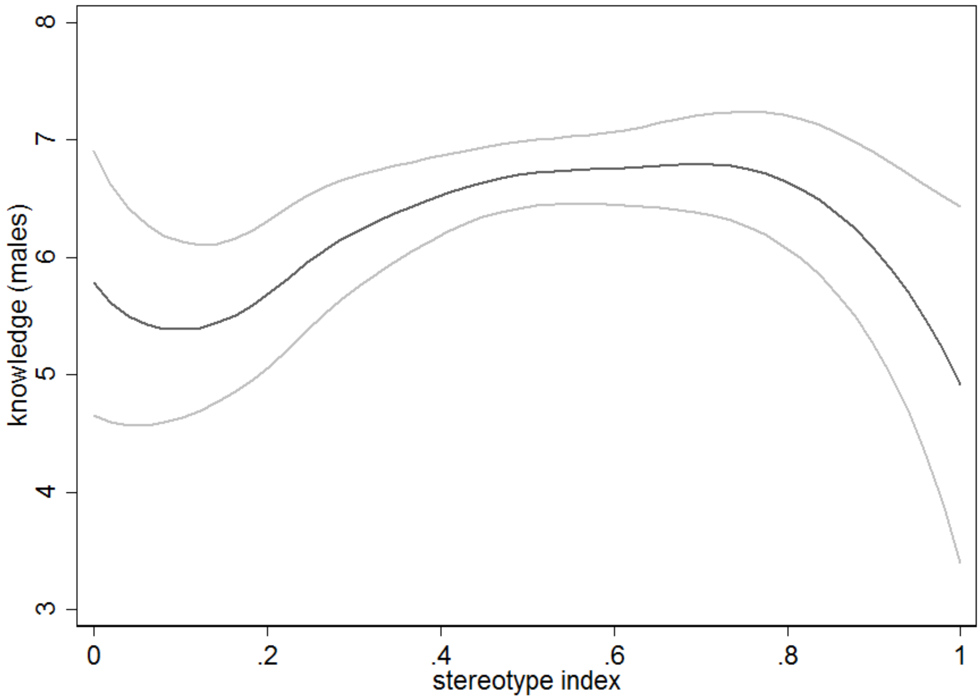
According to our stereotype index, males' beliefs tend to be more biased towards the own gender in all five sub-domains, especially regarding the ability to deal with financial matters and future returns to financial literacy (Table 2.1). Females, in contrast, do not exhibit such self-affirmative beliefs. While they are significantly less biased towards male competency in finance (tests in column 3), both genders believe in a higher male competency in finance.

Following initial data inspection, we allow for a non-linear relationship between the stereotype and financial knowledge measures by using Robinson's semiparametric estimator (Robinson, 1988). Separately by gender, Figure 2.1 shows the nonparametric estimates of the association between stereotype strength and knowledge, controlling for several covariates in the parametric part. Females' financial knowledge tends to deteriorate as the bias in their beliefs increases, while male performance increases with self-affirming belief bias.[6]

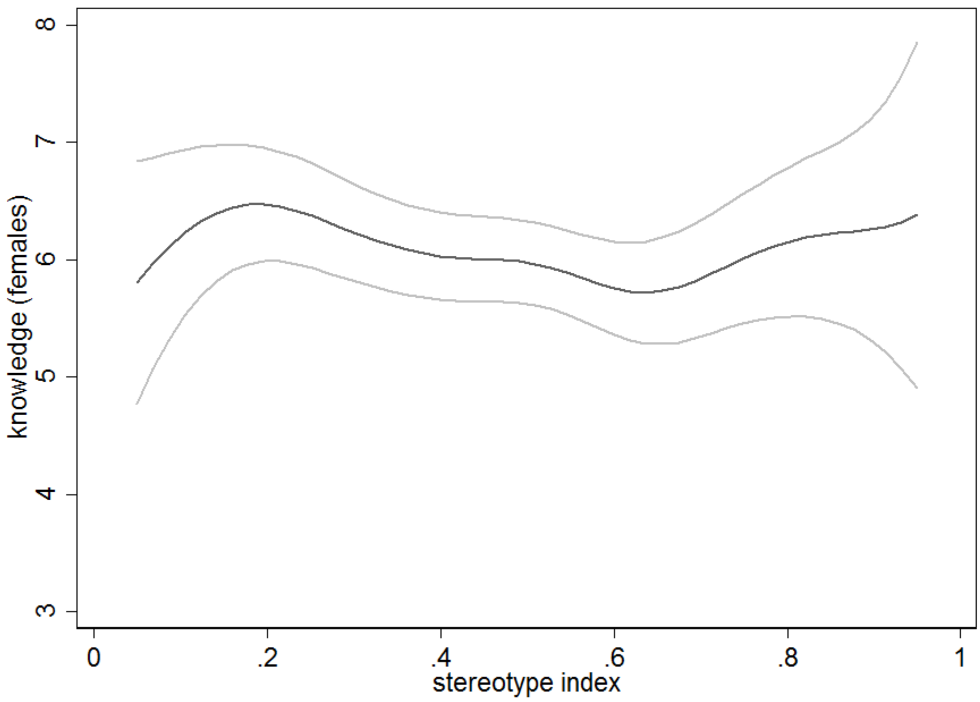
[6]The difference between the nonparametric regression lines for female and male students is more pronounced if we trim the stereotype score at the bottom and top 5% to account for outliers (results available on request).

Figure 2.1: Nonparametric Estimates of the Relationship between Stereotype Index and Financial Knowledge

(a) Males



(b) Females



Note: 95% confidence intervals shown in dotted lines

The last two columns of Table 2.2 report the estimates for the regressors in the parametric part. In contrast to Bucher-Koenen et al. (2014) for adults, we do not find evidence of a link between self-confidence and financial literacy. However, there is a (weakly) positive association of the willingness to take risks and financial knowledge for female teenagers and — as expected — a strongly negative correlation between a low math score and financial knowledge. Students with a low math grade answer, on average, between 0.6 and 1 fewer questions correctly.

2.4 Conclusion

This paper reported data from a field study in German schools among 13 to 15 year old teenagers, confirming the existence of a sizeable gender gap in financial literacy. Differences in numeracy, risk attitudes and self-confidence have been discussed as potential determinants of such differences in prior literature. We do not find systematic differences in these variables by gender. Our data suggest, however, that stereotypical beliefs play a role in the formation of the gender gap in financial literacy. We found no statistically significant knowledge differences between males and females that do *not* display biased beliefs related to financial literacy. For females, financial knowledge deteriorates with stereotype intensity whereas it increases for males.

While we cannot establish a causal relationship between gender stereotypes and financial knowledge, our results show that the gender gap in financial literacy and stereotypes are both present at young ages – consistent with the notion that stereotypes influence investment into financial literacy among teenagers. Further research is needed to establish such causal links and more generally on the formation of gender stereotypes in this and other domains.

Our findings suggest possible ways to improve financial education programs for younger individuals. Such interventions are known to increase self-assessed competence and motivation to engage with finance topics, but not differently so by gender (Lührmann et al., 2015b). Their effectiveness for females might be increased by addressing stereotypes directly. Further, integrating financial education into mandatory school curricula might limit under-investment by groups holding biased beliefs.

CHAPTER 3

Childcare Expansion and Behavioural Health Outcomes: Evidence from Germany

"Early childhood education and care can lay the foundations for later success in life in terms of education, well-being, employability, and social integration, especially for children from disadvantaged backgrounds." *European Commission, Early Childhood Education and Care (2016)*

3.1 Introduction

There is consensus among economists and policymakers that early childhood programs could have beneficial short- and long-term effects (Cunha et al., 2006; World Bank, 2016b). Providing universal childcare is a first-order objective in most developed countries and is frequently put in the center of policy debates and political discussions. The European Council has set clear targets for providing childcare in European Union Member States, where it specifies that "at least 90 per cent of children between 3 years old and the mandatory school age, and at least 33 per cent of children under 3 years old, should have access to formal childcare provision." (Janta, 2013). Yet, more than a decade after these objectives were first outlined, differences in childcare provision within Europe persist. Evidence from the United States is not particularly encouraging either. A study from the U.S. Institute for Early Education Research reports that "only about half of all the three- to four-year-olds are enrolled in preschool" — a figure significantly lower than the OECD average (Economist, 2016). In an effort to increase maternal labour force participation and provide parents with more flexibility, extending childcare is a central topic in education decision making (European Commission, 2011; Janta, 2013). Supporters of early childhood interventions argue that their benefits extend well beyond the children. Better-raised children is equivalent to a lower number of dysfunctional adults at public expense. According to the World Bank (2016b), "every dollar spent on preschool education earns between 6 and 17 dollars of public benefits, in the form of a healthier and more productive workforce".

Empirical research on U.S. childcare programs has documented the large and positive returns of early childcare interventions in later life outcomes (Currie, 2001; Cunha et al., 2006; Ludwig and Miller, 2007; Douglas and Currie, 2011; Heckman et al., 2013). Currie (2001) argues that the main argument for government intervention in early childhood education can be made on the grounds of equity: children from disadvantaged backgrounds start out with lower endowments and are likely to end up with lower allocations. Governments could intervene by spending on childcare programs in order to increase the chances disadvantaged agents have towards a more equal allocation. Douglas and Currie (2011) report positive long-term spillovers from U.S. childcare intervention programs including lower teen pregnancy rates, lower criminal rates and a higher likelihood of obtaining a school leaving certificate. An evaluation of the Head Start program — a program established in 1965 to provide preschool, health and other social services to poor children and their families — by Ludwig and Miller (2007) documents that there is a large drop in mortality rates for children who participated in the program.

Right from the onset, one important distinction has to be made between available evidence from the U.S. preschool programs (e.g. Head Start and Perry School) and evidence from childcare in Europe. The main difference lies on that the U.S. literature has focused on programs that target children from disadvantaged socioeconomic backgrounds (e.g. immigrants, African-Americans), while childcare in Europe is universal, of high-quality and publicly provided to all irrespective of socioeconomic background.

So far, researchers have extensively looked into the effect of universal childcare provision on outcomes such as maternal labour force participation, schooling as well as work-life balance (Felfe et al., 2012; Bauernschuster and Schlotter, 2015; Felfe and Lalive, 2015; Cornelissen et al., 2016). Yet, this paper aims to make a contribution by bringing in evidence on the impact of universal childcare on long-term child health outcomes — where empirical findings remain limited. Recent exceptions include ongoing work by Cornelissen et al. (2016) and van den Berg and Siflinger (2016) who have looked into physical health outcomes in the short run using administrative data. While physical health outcomes are easier to obtain in the short-term, behavioural health outcomes such as the probability of following a healthy diet, the frequency of exercise or the probability of taking up smoking as an adult require the availability of a longitudinal in nature dataset. These types of health outcomes are often of interest to policymakers and economists for two main reasons. First, habits formed early in life could induce higher self-control and subsequently help individuals to stay away from addictive habits such as smoking or excess eating in the future. Second, prevention of these habits could help save on medical expenditure. Cawley and Meyerhoefer (2010) estimate that obesity, which is associated with increased risk of diabetes, heart attack and stroke, accounts for up to 20.6% of U.S. health expenditure.

In this paper, I empirically exploit the nationwide introduction of a legal claim to a place in kindergarten[1] in Germany, first implemented in 1996. This mandated municipalities to provide childcare to children as early as from the age of three. In addition, this entitlement made entry to formal childcare conditional on a date-of-birth cutoff rule. Based on this rule — children who were above 36 months old in August or September of the kindergarten year — were allowed to enter kindergarten in this year at the age of three. Before this rule, most children entered childcare either at the age of four or five.

To the best of my knowledge, I am the first to use this natural experiment to analyse the relationship between eligibility for childcare entry at the age of three and long-term behavioural health outcomes. I focus on Germany for two reasons. First, although Bauernschuster and Schlotter (2015) have looked into the reform in relation to maternal employment, the long-term health implications of this legal entitlement have not been analysed so far. The main motivation behind the entitlement's introduction was to increase enrollment rates for children attending formal childcare below the age of five. In the years that followed, the number of children entering childcare at the age of three or four substantially increased. Data from Cornelissen et al. (2016) show that between 1994 and 2000 the share of children attending public childcare increased from 41.2% to 75.8%. At the same time, the share of children taken care of by their mother at home decreased from 39.3% in 1994 to 18.8% in 2000. Second, universal and highly-subsidized childcare

[1]The official name of the reform in German is *Rechtsanspruch auf einen Kindergartenplatz.* According to the German educational system, Kindergarten is defined as registered childcare services offered to kids between the age of 3 until compulsory school attendance/entry when usually takes place when kids have turned six years old. The children go for at least four hours a day and mostly every day of the week. In Germany, there is a clear distinction between *Kindergarten* and *Kindertagesstatte.* For children who already entered school, the second term denotes extended childcare hours for working parents. In Kindertagesstatte, children stay in school until late afternoon and have lunch within the school facilities.

is available in many other developed countries. Thus, I believe my analysis to be relevant for similar settings outside Germany as well.

Based on findings from the early childhood interventions literature (Carneiro and Heckman, 2003; Cunha et al., 2006), I hypothesize that being eligible for childcare earlier is associated with better long-term behavioural health outcomes such as reduced smoking and better dietary habits. To investigate this claim, I use data from the German Socio-Economic Panel study (SOEP). The SOEP has two main advantages: first, its longitudinal nature allows me to observe individuals' health outcomes as soon as they turn seventeen years old and match their information with household/parental variables. Second, I can observe whether individuals have attended formal childcare at some point in their life, a special characteristic of the SOEP that many administrative datasets do not have.

For the execution of the empirical strategy, I rely on an intention-to-treat design, given that I do not observe the exact age at which children attend childcare for the first time. Hence, I regress health lifestyle outcomes on the eligibility to attend formal childcare at the age of three, according to the cutoff rule. My findings yield no statistically significant effects with respect to health lifestyle outcomes later in life. To address potential violations of the exclusion restriction, I discuss the threats that this cutoff rule introduces for my identification strategy. In particular, how the cutoff rule could affect entrance into elementary school. Next, relying on recent findings from Cornelissen et al. (2016) and van den Berg and Siflinger (2016), I consider the role that family background might play in my setting. Evidence from Felfe and Lalive (2015) and Cornelissen et al. (2016) seems to suggest that there is selection with respect to the type of families that send their children to childcare at an earlier age. The descriptive evidence I have at hand also hints towards this direction. Children coming from better-off families are more likely to be enrolled in childcare earlier (Cornelissen et al., 2016). Therefore, the treatment effect for this selected sample might be negligible, as investments prior to entering childcare might already be sizeable.

The remainder of the paper is organized as follows. Section 3.2 discusses the empirical literature on childcare reforms and long-term outcomes. Section 3.3 outlines the institutional background and section 3.4 explains the data used. Next, section 3.5 presents the empirical strategy and the main findings. Finally, section 3.6 discusses alternative hypotheses before section 3.7 concludes.

3.2 Literature

In this section, I provide an overview of the literature that has looked into childcare reforms in relation to long-term outcomes, ranging from health to social behaviour outcomes.

3.2.1 Childcare Reforms and Long-term Outcomes

There exists a considerable literature that looks into the long-term effects of childcare either with respect to child cognitive outcomes, health and schooling or with respect to work-life balance and maternal labour force participation (Ruhm and Waldfogel, 2012; Black et al., 2014; Bauernschuster and Schlotter, 2015; Cornelissen et al., 2016). Carneiro and Heckman (2003) along with Cunha et al. (2006) were among the first to introduce the hypothesis that early life interventions, including childcare, are crucial for later life outcomes (see also Doyle et al. (2013); Heckman et al. (2013); Elango et al. (2015)). According to Cunha et al. (2006) skill formation is a dynamic process where ability and skills acquired early in life affect later stages. It is argued that, ceteris paribus, a one dollar investment early in life yields higher returns that a one dollar invested at later stages. In support of this hypothesis, there is a growing body of literature studying the importance of early investments for later life outcomes (Case and Paxson, 2008; Currie et al., 2010; Bartling et al., 2012).

The U.S. preschool literature has long identified the long-term gains for participants in Head Start or Perry School (Currie, 2001; Blau and Currie, 2006; Carneiro and Ginja, 2014). Both programs are among the most well-known randomized preschool interventions in the U.S. history. Head Start was established in 1965 and aimed to provide preschool, health and other social services to poor children between the age of three and five and their families (Ludwig and Miller, 2007). The Perry School was introduced a few years later and examined the lives of 123 children, between the age of three and four, born in poverty with increased risk of failing school. The program provided high-quality preschool education and followed individuals until the age of fourty. In this way, researchers were able to measure labour market outcomes such as earnings and the likelihood of holding a job. It has been shown, for instance, that treated children from lower socioeconomic backgrounds were more likely to have higher education attainment and higher earnings in the future compared to their counterparts. Treated children were less likely to commit a crime or to request welfare dependence, i.e. the program provided positive post-intervention social spillovers. Assessment of the effectiveness of universal childcare in Europe, however, remains limited and inconclusive (Havnes and Mogstad, 2011a; Felfe and Lalive, 2015; Baker et al., 2015).

Examples of recent studies, to which my paper relates, look into these issues by exploiting policy reforms that create time and regional variation in the access or the price of childcare. Black et al. (2014) employ a quasi-experimental approach and study subsidies on long-term child outcomes in Norway. The authors conclude that lower childcare prices are associated with an improvement on children's schooling performance as measured by test scores. Another stream has focused on the effects of center-based childcare on school readiness, where Felfe and Lalive (2013) find significant positive effects before the age of three. Other studies, including that of Kottelenberg and Lehrer (2013), highlight the importance of looking into treatment effect heterogeneity in order to better understand which groups of children could benefit most from childcare interventions.

My paper closely relates to two recent studies that look at the interplay between childcare and health outcomes. Ongoing work by van den Berg and Siflinger (2016) uses Swedish administrative data to evaluate the effect of an exogenous shock to childcare on medium term child health outcomes. The authors show that higher priced childcare has an effect across different age groups. For kids up to the age of five, there is an improvement in physical health, whereas children between the age of six and ten exhibit better mental and psychological functioning and hence less behavioural or developmental complications. On another study, Felfe and Lalive (2015) use administrative data from a large German state to find positive health effects for students who come from a migrant or disadvantaged background. Combined with a marginal treatment effects analysis, a 'modest expansion' of childcare positively contributes towards children's developmental improvement (Felfe and Lalive, 2015).[2] When it comes to extending schooling, Grossman and Kaestner (1997) and Lleras-Muney (2005) show that health information and better decision making act as channels for the positive relationship between education and health. In another study, DeCicca et al. (2008) employ an instrumental variable approach exploiting variation by

[2]A handful of studies has looked into the effect of extending kindergarten eligibility for each gender. Baker et al. (2015) support that the effect on non-cognitive skills may differ between males and females over the long run. Felfe and Lalive (2015), Datta Gupta and Simonsen (2015) and Kottelenberg and Lehrer (2016) move along similar lines.

U.S. state education policy to find that graduating from high school is associated with less smoking. Estimates are less precise when looking into obesity.

3.3 Institutional Background

In this section, I provide a detailed description of the institutional background that accompanied the legal claim to a place in kindergarten, including the type of childcare that the entitlement targeted and the population affected. Eventually, I use the policy reform as a means to empirically evaluate the effect of being eligible for earlier entry into formal childcare[3] on later health outcomes.

3.3.1 The German Childcare System

Curriculum & Quality Standards

Germany's childcare system, similar to that of other European countries, is highly regulated and governed by a set of clearly defined quality standards. First, districts and municipalities have little space for making changes or additions to the standard curriculum. This is usually because the curriculum is set under scrutiny and its structure is determined at the federal level. Among the activities included in the curriculum encourage teachers to support children in engaging in playful activities such as circle play, painting and reading or physical activities (Felfe and Lalive, 2015). Playground rules also follow a protocol, namely that a group of ten children should always be under the supervision of at least one certified education personnel and one assistant. Furthermore, it is also important that children start building their first social interactions, either with other peers or with staff in the childcare center. Overall, a typical day at the kindergarten is a combination of informal learning and various games. In essence, the main goal of the federal curriculum is to enhance day-to-day social interaction among children and between children and teachers (Cornelissen et al., 2016).

Second, quality standards include guidelines with respect to teachers' former education and the student-teacher ratio. Teachers are expected to have completed a two-year state-certified vocational training, followed by a one-year internship as a childcare teacher. Concerning the student-teacher ratio, it is regulated and it should not exceed twenty-five

[3]Throughout the paper, earlier entry into childcare is equivalent to children being eligible to enter childcare at the age of three, following the eligibility cutoff rule. The two terms are used interchangeably.

children per two teachers. In a more general assessment, Germany's student-teacher's ratio performs rather average when compared to other OECD countries.[4] (Organisation for Economic Co-operation and Development, 2014; Cornelissen et al., 2016).

Types of Care

There is formal and informal childcare. The types of formal childcare vary from half-day care (8am - 12am) to full-day care (8am - 3pm). Statistical data from the Deutsches Statistisches Bundesamt (2016) indicate that the majority of children receiving formal childcare attend part time care, i.e. four hours in the morning. Children, who do not attend formal childcare, usually stay at home to receive informal care by a caregiver, such as the mother or the grandmother. As table 3.1 informs, before the introduction of the reform, about 41% of children were enrolled in public childcare while this number had jumped to 75.8% by 2000. On the contrary, the figures for informal family (e.g. parents/relatives) and exclusive maternal care are particularly high in 1994 (58.3% and 39.3% respectively) but drop considerably by the beginning of the 21st century (see table 3.1).

State Subsidization

Childcare in Germany heavily relies on state subsidization. The reasons for this highly subsidized market are the large administrative and investment costs associated with setting up childcare facilities. Illustratively, recent work by Felfe and Lalive (2015) shows that — of the total operating cost of 14.1 billion Euros in 2006 — state subsidies covered about 79% while the remaining was covered by parents and other private organizations by 14% and 7% respectively. Fees range from 0 to 600 Euros per month, depending on income and family size. Larger families pay less while better-off families pay more. In general, due to the heavily subsidized setting, the monthly price that parents are asked to pay is considerably below the marginal cost (for details, see Felfe and Lalive (2015); Felfe and Zierow (2015)).

3.3.2 Legal Claim to a Place in Kindergarten

In January 1996, Germany introduced a legal claim to a place in kindergarten according to which every child was eligible to attend formal, center-based childcare from age

[4]Germany's 12:5 student-teacher ratio lies between the 8:1 ratio applicable to three- to seven-year-olds in the United Kingdom's center-based programs and that of 25:1 applicable to French childcare programs.

three until school entrance. This legal entitlement applied to public, highly subsidized and half-day care across all federal states and municipalities. Following its introduction, there was a substantial increase in the demand for a spot in formal facilities. As the supply of formal childcare did not respond proportionally, many municipalities were unable to provide enough places that would accommodate requests from parents. Due to organizational frictions and excess demand, implementation proved to be more troublesome than expected. As a result, the German Federal Parliament, *Deutscher Bundestag*, introduced another legislative initiative that aimed to better control for the number of children entering childcare at the age of three.

This initiative resulted in the adoption of an eligibility cutoff rule – right after the 1996 legal claim – in an effort to control increased demand for an available place in formal childcare. Municipalities were obliged to secure a place **only** to eligible children, i.e. children who were older than 36 months and less than 48 months at the time of the start of the kindergarten year. In the German educational system, the start of the kindergarten year usually coincides with the start of the school year, namely either in August or in September depending on the state. By definition, eligible were the children who entered formal childcare following their third birthday (i.e. 36 months old). If a child was below 36 months old, it would have to wait for one more year to enter formal childcare. Politically, this legislative initiative managed to take off part of the pressure lying on municipalities' shoulders with respect to how they would allocate the excess demand created after the 1996 legal claim. Although the eligibility cutoff date-of-birth rule was introduced as a temporary measure aimed to last until the end of 1998, its use continued long after.[5]

Using data from the Micro Census, Bauernschuster and Schlotter (2015) show that childcare rates for three and four year old children increase considerably after the 1996 claim. The attendance of three-year-olds in particular, increased by more than 25 percentage points between 1996 and 2003. Concerns may arise with regards to the type of childcare institution (e.g. public, private, church-provided) that were driving this increase. Although the SOEP data do not allow me to further disentangle this number, related work (Felfe and Zierow, 2015; Felfe and Lalive, 2015) reports that the share of private childcare facilities in the 1990s was negligible.

[5] In my main sample, I test the hypothesis only for cohorts born between 1992 and 1997 when I expect my instrument's relevance to be strongest

Informal Childcare: Pre & Post Reform

Using the SOEP data, I further look into the extent to which the introduction of the 1996 reform might have led to changes in informal childcare, i.e. whether after the reform children that used to be taken care by mothers/grandmothers are now receiving formal care. Prior research looking into childcare expansion and maternal employment from Havnes and Mogstad (2011a) and Bauernschuster and Schlotter (2015) seems to hint towards this direction. For instance, Havnes and Mogstad (2011a) analyse the effects of a staged expansion of subsidized childcare in Norway to find that instead of an increase in maternal employment subsidies crowd out informal childcare arrangements.

Table 3.1 and table 3.2 illustrate childcare arrangements for 3-year-olds in West Germany. Evidence from Cornelissen et al. (2016) in table 3.1 shows that the alternative to formal childcare is almost exclusively informal care by either parents or grandparents. The authors use data from the German Family Survey to illustrate the large shift from informal childcare (mainly that provided by mothers and relatives) to public childcare. For instance, while exclusive maternal care corresponds to 39.3% of total childcare arrangements in 1994, this number drops to 18.8% by 2000. On the contrary, the share of 3-year-olds attending public childcare, both of working and non-working mothers, almost doubles between 1994 and 2000. Table 3.1 thus provides a bigger picture explanation of the evolution of childcare arrangement for 3-year-olds shortly before and a few years after the 1996 reform.

Table 3.2 uses the SOEP sample to perform a robustness type of check between a sample of children not affected by the reform – born between 1987-1991 – and a group of children born between 1992 and 1997 when childcare entry followed the cutoff date-of-birth rule. I further split these into subgroups based on eligibility for childcare entry at the age of three following the cutoff rule. Formal childcare includes heavily subsidized nurseries and public childcare (*kindergartens*), operating either full- or part-time. Informal care includes children who are cared by relatives, friends or other types of unlicensed caregivers. Comparing columns 1 and 2, the share of children receiving formal childcare increases after the introduction of the reform, while informal childcare decreases. The change is mainly coming from *care by relatives* and *care by no-one*.

Table 3.1: Childcare Arrangements for 3-year-olds — Pre & Post 1996 Reform

Type of childcare	1994 Share	2000 Share
Public childcare	41.2%	75.8%
Only family care (parents/relatives)	58.3%	22.7%
Exclusively maternal care	39.3%	18.8%
Informal care (nanny, other non-relatives)	1.2%	1.5%
Maternal labour force participation (3-year-olds)	31.2%	38.7%
Public childcare, children of working mothers	42.9%	81.5%
Public childcare, children of non-working mothers	40.5%	72.2%

Notes: The figure is adopted from Cornelissen et al. (2016). Unfortunately, the SOEP dataset does not allow me to obtain detailed data on the type of the childcare arrangements for 3-year-olds before and after the reform. The table provides, for the years 1994 and 2000, information on childcare arrangements distinguishing between public childcare, only family care by parents or close relatives, only maternal care and care by a child-minder or nanny. The table also reports maternal labour force participation rates for mothers of 3-year-olds along with public childcare attendance rates of 3-year-olds. The data are based on Cornelissen et al. (2016) calculations along with data from the Family Survey from the German Youth Institute (DJI), Munich, 2nd and 3rd wave (1994 and 2000). The sample refers to 3-year-olds in West Germany and consists of 262 children in 1994 and 354 children in 2000.

Table 3.2: Childcare Trends — Pre & Post 1996 Reform

	Pre-Reform (placebo) Eligible	**Post-Reform** Eligible
Formal Childcare		
Nursery care	<1%	<1%
Kindergarten	30%	33%
School	29%	32%
Daytime care	26%	28%
Informal Childcare		
Care by relatives	<2%	<1%
Care by friends	<1%	<1%
Care by paid sitter	<1%	<1%
Care by no-one	12%	5%

Notes: The pre-reform/placebo group replicates a *pseudo* eligibility cutoff rule and consists of all children born between 1987 and 1991 who were not affected by the reform at any time of kindergarten entry. The post-reform group includes children born between 1992 and 1997 who were actually eligible for entry into formal childcare at the age of three. The two columns report the average share of children attending each form of childcare at the given point in time. Data: SOEP, own calculations.

3.4 Data

In this section, I use individual level health and childcare attendance data from the German Socio-Economic Panel study (SOEP) together with information on the timing of childcare entry across federal states and years. For further details on how the data were constructued, I have included a more detailed description in Appendix B.2.

3.4.1 The German Socio-Economic Panel

I use rich, individual level data taken from the German Socio-Economic Panel (SOEP), a representative and longitudinal household survey in Germany. The same private households, individuals and families are tracked along the years, dating back to the start of the first SOEP survey wave in 1984. The data provide a range of information including health outcomes, living conditions of individuals, demographics, employment and schooling variables as well as the relationship between inter-dependencies among all these areas. Particularly interesting for my empirical approach is that children who were treated while in formal childcare are observed and interviewed individually once they turn seventeen years old. To gain statistical power, I take into account the full SOEP sample which includes both West and East Germany, although I expect the treatment effect to be weaker in the East. East Germany had traditionally a higher share of mothers in the labour force. Before the unification with the West, it had already established a formal childcare support system that offered higher degrees of flexibility for working mothers. Formal childcare support was also often available to children below the age of three. This system continued to exist even after the early 1990s. Yet, this might not exclude the possibility of certain municipalities and schools in the East to have taken advantage of the eligibility cutoff rule in order to better deal with childcare place scarcity. In West Germany, on the contrary, childcare for children below the age of three was not in place until many years later, while attendance for children between the age of three and six (i.e. until school entry) was still limited. This was especially the case for children between three and four years old.

3.4.2 Behavioural Health Outcomes

The set of health outcomes that I am investigating constitute behavioural/lifestyle type of outcomes observed in the dataset after the individual has turned 17 years old.[6] These include: a) a dummy variable on whether the individual follows a healthy diet, b) a dummy variable on whether s/he smokes, c) a continuous measure of obesity constructed using each individual's Body-Mass-Index (BMI)[7] and d) a dummy variable for the frequency of sports. Obesity and smoking have been consistently ranked to have harmful short- and long-term effects. The World Health Organization (2016a) stresses that while smoking's short-term consequences involve milder respiratory and non-respiratory diseases, in the long run lung cancer could also be a threat. Obesity is another cause of concern, as being overweight is associated to major risk factors including diabetes and cardiovascular diseases. From a policy point of view, the costs that the health care system incurs to tackle such diseases and their negative spillovers are rather high (Cawley and Meyerhoefer, 2010). Therefore, the returns from following a healthier diet and increasing exercise in early adulthood could lead to some type of habit formation that transmits into a healthier adult lifestyle overall.

3.4.3 Childcare Attendance and Demographics

A novel feature of this dataset is that I can directly link health lifestyle outcomes later in life with formal childcare attendance observed earlier in life. For every individual in my dataset, I construct a binary indicator that takes the value of one if the individual has attended childcare at any point in his/her life while the variable is assigned a value of zero if this is not the case. However, I am not in the position to observe whether the individual attended childcare exactly at the age of three. This is because while for some individuals the survey year coincides with the year they started attending care, for others it is a regressive question asked through the household questionnaire at a later stage in life. Hence, it is not possible to observe for the full sample whether individuals started attending formal childcare earlier, later or exactly at age three. If I only tracked those

[6]The terms behavioural health outcomes or health lifestyle outcomes are used interchangeably. The SOEP first performs personal interviews when the individual has reached the age of 17. In my sample, this is the age of the first recorded health outcomes' observations.

[7]Body Mass Index calculated as weight(kg)/[height(m)2]

individuals in the cases that this was possible, the sample size would shrink by a sizeable magnitude.

Rich demographic and parental information allow me to additionally control for a set of observed differences between treated and untreated population. These controls include binary indicators for: child's gender – assigned the value of one if male – child/mother nationality – assigned a one for German – maternal high school diploma and an employment indicator – assigned the value of one if the mother is working full-time. Continuous measures on child/mother age, mother's years of schooling as well as household size and number of siblings are further included. Combining this information with federal state — *Bundesländer* — data on the timing of childcare entry across federal states (Deutsches Statistisches Bundesamt, 2016), I am able to identify whether the child was above or below the eligibility cutoff date-of-birth rule at the time of the start of childcare.[8] My sample consists of all those individuals whose health lifestyle outcomes later in life can be observed, i.e. they are not missing. This implies that my sample is reduced to about 670 individuals and mother-child pairs born between 1992 and 1997. For the first years of their life, individuals can be followed through the household ID but once they reach adulthood I can observe them through a personal ID. In addition, I include a sample of about 1600 children born between 1987 and 1991 to be used as robustness checks. Children born before 1992 have been tracked for a longer time by the SOEP and hence the number of observations available for their health outcomes is higher. This explains why the two samples differ in terms of observations.

Table 3.3 illustrates summary statistics by eligibility group in the sample. Column 1 of table 3.3 lists children who were less than 36 months old at the time of the childcare start and hence non-eligible for early childcare entry (control group). I compare these statistics to column 2 where the treated group is shown, i.e. children who were older than 36 months at the time of formal childcare start (August or September). In column 3, I report the p-values of their difference after performing a t-test. To avoid that these

[8]The first stage analysis of section 3.5 requires detailed information on the time of the start of kindergarten year as it varies by state. Combining the SOEP data on German states with information from the *Kulturministerkonferenz* aligned with Bauernschuster and Schlotter (2015), I constructed the *Einschulung* month for each survey year at each state. The use of survey year and household specific identifiers lies on the fact that school entrance dates vary within states and across years. This implies that if the eligibility rule is strictly determined by the month of the kindergarten entrance, it is important to know what date for each cohort interviewed in my sample.

Table 3.3: Summary Statistics — by Eligibility

	Age at time of childcare entry		Difference
	Control Group	Treatment Group	Test $H_0 : x_c = x_t$ (p-value)
	(1)	(2)	(3)
Children's Demographics			
Child's age at childcare entry	30.069	42.255	0.000***
	(3.312)	(3.998)	
Gender (= Male)	0.531	0.471	0.565
	(0.499)	(0.499)	
Nationality (= German)	0.916	0.913	0.003***
	(0.077)	(0.080)	
Household Size	4.119	4.215	0.123
	(1.085)	(0.979)	
No. of siblings	2.644	2.709	0.459
	(1.003)	(1.008)	
Mother's Demographics			
Mother's Age	36.548	36.693	0.753
	(8.567)	(9.064)	
Nationality (= German)	0.877	0.879	0.088*
	(0.329)	(0.326)	
High School Diploma	0.653	0.688	0.177
	(0.477)	(0.464)	
Years of Schooling	11.837	11.837	0.369
	(2.804)	(2.493)	
Employment Indicator	0.403	0.421	0.156
	(0.491)	(0.494)	
N	303	376	679

Notes: The sample consists of all children and all mothers-children pairs whose health outcomes are observed later in life. Children in the sample are born between years 1992 and 1997. Age is reported in months. Column (1) reports children in the control group, i.e. between 24 and 35 months at the time of childcare entry. Column (2) reports children in the treatment group, i.e. between 36 and 48 months at the time of entry. Column (3) performs a t-test to check whether the difference between columns (1) and (2) is statistically significant. p-values are reported in column(3). The table shows means; standard deviations are given below in parentheses. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

differences drive my results, I include child as well as parental background controls in all my regressions. Age differences are significant by construction.

Looking at the nationality indicator, it is evident that treated children and their mothers are significantly more likely to hold a German nationality (see table 3.3). This of course raises concerns of strategic planning of births from the mother's side. It could be that German mothers have higher chances to know about the reform and therefore plan their time of birth to gain access to childcare earlier. Despite controlling for nationality later in regressions, this gives rise to concerns about a selected sample towards children from a better socioeconomic background, as indicated by Cornelissen et al. (2016). To check if this is also the case in the pre-reform (placebo) sample, I look into the nationality statistics of children who attend / do not attend formal childcare and are born before 1992. I find that birth dates are systematically different by nationality. Children who attend childcare and are born before 1992 are 3 percentage points more likely to hold a German nationality compared to children who do not attend childcare and are also born before 1992. This is a statistically significant difference (p-value=0.001). Regarding gender, household size, number of silbings, mother's age as well as mother's education and employment level, they are not significantly different for the two groups at conventional significance levels.

3.5 Empirical Analysis

In this section, I first look into an ordinary least squares setting and explain why results using this approach might be biased. I move on to an intention-to-treat (ITT) design where I regress health outcomes later in life on the eligibility to attend childcare at the age of three. In my main empirical specification, eligibility is constructed based on a date-of-birth cutoff rule introduced directly after the 1996 legal claim to a place in kindergarten. To pin down this relationship, I follow closely the empirical approach in Bauernschuster and Schlotter (2015).[9] Finally, I discuss the paper's results and I address possible threats to the identification of the findings.

3.5.1 OLS Estimates

Ideally, I would like to look into the intensive margin of extending childcare attendance on health outcomes. In particular, I would like to analyse whether one more year of childcare attendance (starting at age three instead of four) is positively associated with long-term behavioural health outcomes. Yet, the structure of the data (e.g. exact age of childcare

[9]Similarly to Angrist and Krueger (1991) where their quarter-of-birth strategy is followed by the fact that most states in the United States allow children to enter school in the calendar year when they turn six years old. As in my setting, school start age is a determined by the date-of-birth.

entry unobserved) only allows me to investigate whether higher childcare attendance in general improves health outcomes later in life. In what follows, I look into an ordinary least squares model. In the absence of any endogeneity, the standard econometric model for the long-term effect of attending formal childcare on behavioural health outcomes could be expressed as shown in the following equation:

$$Health_i = \alpha_s + \tau_t + \phi ChildCareAttendance_i + \beta \cdot X_i + \epsilon_i \tag{3.1}$$

$Health_i$ is defined as individual *i's* health lifestyle set of outcomes when s/he has turned 17 years old. These four lifestyle outcomes are measured by: a) a binary indicator for following a healthy diet, b) a binary indicator for smoking, c) a continuous measure for Body-Mass-Index (BMI) and d) a binary indicator for the frequency of sports. X_i is a vector of sociodemographic characteristics available at the individual level. These explanatory variables include family characteristics (years of schooling, high school graduation, full employment indicator, income, nationality and age) along with child demographics such as nationality, age, gender, number of siblings and household size. $ChildCareAttendance_i$ is a dummy variable denoting whether or not the individual has attended formal childcare at any point during his/her life. α_s denotes federal state controls and accounts for unobserved, time-constant differences across the sixteen federal states. τ_t are survey year controls accounting for time trends. Finally, the error term, ϵ_i, contains all unobservable influences on the set of health outcomes. In particular, it contains the effects of unobservable individual and family characteristics.

The main disadvantage in applying ordinary least squares is that OLS gives rise to biased estimates if unobserved characteristics and $ChildCareAttendance_i$ are correlated. Measurement error in childcare attendance would lead to a downward bias which would in turn underestimate the effect of attending childcare on children's future health outcomes. Another caveat is that individuals may select into childcare with respect to unobserved factors, such as student's ability and family background characteristics. For example, families where both parents work full-time and face time-constraints may insist on their children getting a place in a childcare facility. Furthermore, as Bauernschuster and Schlotter (2015) have shown, earlier childcare attendance is associated with a higher probability of maternal employment. Hence, attending childcare may positively impact health lifestyle

Table 3.4: OLS Regressions — Standard Specification

Dep. var.: Health outcomes (Age 17)	Healthy Diet (1)	Smoking (2)	BMI (3)	Sports (4)
Childcare attendance	-0.1316*	0.0637	-0.9545	0.0455
	(0.0717)	(0.0695)	(0.5595)	(0.0784)
Year controls	Yes	Yes	Yes	Yes
Federal state controls	Yes	Yes	Yes	Yes
Individual level controls	Yes	Yes	Yes	Yes
Observations	556	556	556	556
R-squared	0.17	0.14	0.14	0.19

Notes: The table reports OLS regressions. Health outcomes are measured around the age of 17, i.e. as soon as the individual is an adult. Outcome and control variables are defined in section 3.4. The sample includes all children born between 1992 and 1997. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

outcomes later in life not merely through attendance but also via better education tracks and/or family income.

Table 3.4 reports baseline OLS regressions for the four health lifestyle outcomes, namely: a binary indicator for following a healthy diet, an indicator for smoking, a measure of Body-Mass-Index (BMI) and an indicator for the frequency of sport activity. Columns (1) to (4) account for a set of sociodemographic variables (child's nationality, gender and age, household characteristics and mother's education, employment and income) along with federal state and year controls. Findings are noisy and coefficients echo instability with respect to the direction of the effect. For instance, childcare attendance seems to be associated with a significant decrease in the probability that the individual follows a healthy diet, while it is positively correlated with the probability of smoking – ceteris paribus. Findings indicate that endogeneity of childcare attendance might bias the effect on long-term health outcomes. For this reason, I look into an intention-to-treat approach next.

3.5.2 Intention-to-Treat Approach

OLS fails to solve the endogeneity problems between childcare attendance and health outcomes observed later in life. I proceed to apply an intention-to-treat (ITT) design that looks at the reduced-form effect of being eligible for childcare entry at the age of three on long-term behavioural health outcomes. In the section that follows, I initially carry out a first stage approach where I confirm the instrument's relevance. I then present my reduced-form results.

First stage

To regress childcare attendance on being eligible for entry into childcare at the age of three, I construct a binary indicator – namely – $EarlierEligibility_i$. The sample includes all individuals born between 1992 and 1997 in Germany. This indicator is constructed symmetrically around the cutoff age of 36 months. $EarlierEligibility_i$ takes the value of one for a child who is older than (or equal to) 36 months but less than (or equal to) 48 months at the time of kindergarten start. It is assigned a zero if the child is older than (or equal to) 24 months but less than (or equal to) 35 months at the time of kindergarten entry. The time of kindergarten start is precisely defined as the month when kindergarten starts in Germany. This varies across states but in all cases in my sample, it is either sometime in August or on some date in September.[10] The date-of-birth cutoff rule creates variation in childcare attendance that should be exogenous to confounding factors determining childcare and behavioural health outcomes simultaneously.

Understanding the eligibility cutoff rule is crucial. Therefore, the following example aims to better illustrate the rule's underlying principle. Assume one looks at September 1996 as the kindergarten entry month. This date coincides with the first year when the reform was introduced. If I assume that kindergarten in a federal state, F , starts in September (of 1996), then there are two scenarios. The first scenario is that Child A is born in September 1993. Child A is 36 months old at the time when kindergarten starts in federal state F. Hence, according to the eligibility cutoff rule, Child A is eligible to attend formal childcare at the age of three. The binary variable, $EarlierEligibility_i$ takes the value of one as Child A belongs to the treatment group. This is the case for all children born in September or the months **before** of the same year. The second scenario assumes that Child B is born in October 1993. Child B is 35 months old at the time when kindergarten starts in federal state F. Hence, following the eligibility cutoff rule, Child B is not eligible to attend formal childcare at the age of three and has to wait for one more year. The binary indicator, $EarlierEligibility_i$, takes the value of zero. Child B belongs to the control group and this is the case for all children born in October or **later** of the same year.

[10]For details on the construction of the time of kindergarten start by state, please refer to section 3.4.

The equation below illustrates the regression:

$$ChildCareAttendance_i = \alpha_s + \tau_t + \gamma EarlierEligibility_i + \delta \cdot X_i + \epsilon_i \qquad (3.2)$$

$ChildCareAttendance_i$ is a binary variable that is either assigned a zero or a one depending on whether the individual in the SOEP dataset has attended childcare at any point during his/her lifetime. Observing childcare attendance is an important advantage of the SOEP. $EarlierEligibility_i$ has been already defined as a dummy variable that takes the value of one for children older than (or equal to) 36 months but less than (or equal to) 48 months at the time of kindergarten start. These children belong to the treatment group. The control group includes children who are older than (or equal to) 24 months but less than (or equal to) 35 months at the time of entry.

α_s denotes federal state controls and accounts for unobserved, time-constant differences across the sixteen federal states. In addition, the fixed effects account for unobserved differences between East and West Germany that remain constant over time. τ_t are survey year controls accounting for time trends. X_i is a vector of individual controls including child's age, household size, nationality of the child as well as mother's age, nationality, education level and employment indicator. Finally, as usually, ε_i is a mean zero error term that by assumption is uncorrelated with $EarlierEligibility_i$. Standard errors are clustered at the individual level. To increase confidence in my empirical strategy, I carry out robustness checks. In particular, I construct a placebo sample and a pseudo childcare entry cutoff for cohorts born before 1992 and hence they were not affected by the legal entitlement.

Table 3.5 shows the first stage regressions. In particular, column (1) reports the effect of $ChildCareAttendance_i$ on $EarlierEligibility_i$ in the absence of any individual level or federal state/year controls. The relationship is positive and statistically significant. Yet, an obvious caveat is that this effect may be mechanically driven when controls are not included. Next, I include federal state/year controls and individual level controls which do not affect the precision of the estimated coefficients (see columns (2) and (3) in table 3.5). This confirms that the instrument is a relevant predictor of childcare attendance. When looking at the full specification of column (3), the effect seems to remain almost unchanged by the inclusion of child's age — among other individual level controls and

Table 3.5: First Stage — Standard Specification

Dep. var.: Childcare Attendance	Full sample (1)	Full sample (2)	Full sample (3)
Earlier childcare eligibility	0.1334**	0.1484**	0.1648**
	(0.0622)	(0.0526)	(0.0582)
Year controls	No	Yes	Yes
Federal state controls	No	Yes	Yes
Individual level controls	No	No	Yes
First stage F-test	3.38	7.98	6.19
Observations	678	678	544
R-squared	0.01	0.11	0.15

Notes: The table reports baseline first stage regressions. The dependent variable is actual childcare attendance, a dummy taking the value of one if the individual in my sample has attended formal childcare at any point during the period of observation. $EarlierEligibility_i$ – the explanatory variable – is a dummy taking the value of one if individual was between 36 and 48 months old at the time of childcare entry. The sample in columns (1) to (3) includes all children born between 1992 and 1997. Individual controls are defined in section 3.4. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

fixed effects — even if the F-test is weaker than in column (2). In particular, being eligible for earlier entry into childcare increases the probability of actual childcare attendance by around 16.5 percentage points – ceteris paribus. This raises the probability of attending childcare from 11.7 to 28.2 percentage points when relating this coefficient to the mean value of $ChildCareAttendance_i$ before the introduction of the 1996 legal claim (i.e. before the reform).

Table 3.6 extends the first stage regressions and performs a set of robustness checks. In particular, I look into three different samples. First, I break down my full sample to look into West and East Germany separately. East Germany had traditionally a higher share of mothers in the labour force. Before the unification with the West, it had already established a formal childcare support system that offered higher degrees of flexibility for working mothers. Formal childcare support was also often available to children below the age of three. This system continued to exist even after the early 1990s. Yet, this might not exclude the possibility of certain municipalities and schools in the East to have taken advantage of the eligibility cutoff rule in order to better deal with childcare place scarcity. In West Germany, on the contrary, childcare for children below the age of three was not in place until many years later, while attendance for children between the age of three and six (i.e. until school entry) was still limited. This was especially the case for children

Table 3.6: First Stage — Placebo Regressions

Dep. var.: Childcare Attendance	West Germany (1)	East Germany (2)	Placebo (1987-1991) (3)
Earlier childcare eligibility	0.1189*	0.0412	-0.0388
	(0.0609)	(0.0607)	(0.0254)
Year controls	Yes	Yes	Yes
Federal state controls	Yes	Yes	Yes
Individual level controls	Yes	Yes	Yes
Observations	353	178	1565
R-squared	0.27	0.45	0.10

Notes: The table reports a set of robustness checks using three difference samples. As before, the dependent and explanatory variables are $ChildCareAttendance_i$ and $EarlierEligibility_i$ respectively. Column (1) reports regression outcomes using the West Germany sample. Column (2) only uses observations from East Germany while column (3) reports findings from the placebo check. The placebo sample includes all children born between 1987 and 1991, i.e. children not affected by the 1996 reform at any stage of childcare entrance. Columns (1) to (3) include a full set of individual and federal state/year controls. Individual level controls are defined in section 3.4. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

between three and four years old (for details, see section 3.4). Therefore, in table 3.6, I would expect to see a strong and significant effect for the West Germany sample (column (1)) where the reform was introduced for the first time in 1996, while weaker should be the effect for East Germany (column (2)). As expected, the coefficient in column (1) goes towards the right direction. Being eligible for entry into childcare at the age of three in West Germany increases the probability of actual childcare attendance by 11.9 percentage points conditional on all other variables being held constant. Looking at the sample of East Germany, in column (2), the effect is positive but not statistically significant.

In addition, I look into a placebo sample which only includes children born between 1987 and 1991 (see column (3) of table 3.6). Given that the reform was first introduced in 1996 and affected cohorts born only after 1992, I would expect to see no positive correlation between $EarlierEligibility_i$ and $ChildCareAttendance_i$ before the legal claim was put in place. Indeed, if anything, the coefficient goes towards the opposite direction and is insignificant. In a next step, I plot histograms for every cohort at the 'age at entry' in order to check for bunching around the cutoff dates (histograms not reported). The main concern might be that families sending their children to formal childcare — the years following 1996 — could have strategically planned the birth date of their child to comply with the eligibility rules. The findings do not raise suspicion with respect to bunching after the first year of the reform.[11]

Intention-to-Treat Results

For reasons discussed later, a two-stage-least squares approach may fail to satisfy the assumption that childcare attendance has an effect on future health outcomes only through the earlier eligibility channel. I therefore turn to an intention-to-treat setting, where I regress a set of health lifestyle outcomes on $EarlierEligibility_i$. Following Carneiro and Heckman (2003) and Douglas and Currie (2011), I expect the direction of the effect to be positive for those dependent variables related to a healthy diet and to sports, while it should be negative for those related to smoking and BMI. The following equation expresses the relationship at hand:

$$Y_i = \alpha_s + \tau_t + \psi EarlierEligibility_i + \theta \cdot X_i + e_i \tag{3.3}$$

Y_i is a set of individual health lifestyle outcomes first measured when the individual turns seventeen years old. $EarlierEligibility_i$ is the same binary indicator that takes the value of one for individuals whose age (in months) is >= 36 months and <= 48 months at the time of childcare entry. The coefficient of interest is ψ: the effect of the eligibility cutoff on a set of health outcomes. α_s accounts for unobserved, time-constant differences across federal states while τ_t are survey/year controls accounting for time trends. The standard set of sociodemographic and family controls are included in $\theta \cdot X_i$ and e_i is a mean zero error term. Standard errors are clustered at the individual level.

In table 3.7, I estimate the intention-to-treat (ITT) effects by regressing health outcomes on my eligibility cutoff instrument that induces entry into formal childcare at the age of three. To account for the fact that statistically significant differences between the control and the treatment group might drive my results, I include a set of sociodemographic controls (gender, nationality, household size and maternal characteristics) along with federal state and year dummies in all regressions. My findings in table 3.7 are rather small in magnitude and they suggest that there is not a statistically significant relationship between being eligible for one more year in childcare and behavioural health outcomes

[11] I also looked into the full sample first stage regressions by cohort and by year. The first year of the reform exhibits the strongest pattern. Ceteris paribus, this translates to an increase in childcare attendance by 30 percentage points due to eligibility, significant at 1%. For subsequent years, the direction of the coefficient remains positive but the magnitude of the effect weakens.

Table 3.7: Baseline Outcomes: Intention-to-Treat

Dep. var.: Health outcome (Age 17)	Healthy Diet (1)	Smoking (2)	BMI (3)	Sports (4)
Earlier childcare eligibility	0.0013 (0.0536)	0.0693 (0.0594)	0.1473 (0.4548)	0.0562 (0.0563)
Year controls	Yes	Yes	Yes	Yes
Federal state controls	Yes	Yes	Yes	Yes
Individual level controls	Yes	Yes	Yes	Yes
Observations	556	556	556	556
R-squared	0.16	0.14	0.14	0.19

Notes: The table shows reduced-form estimates. The dependent variables are a set of health outcomes that include: a binary indicator for following a healthy diet, a binary indicator for smoking, a continuous measure for Body-Mass-Index (BMI) and a binary indicator for the frequency of sports. $EarlierEligibility_i$ is the main explanatory variable. The first stage is reported in table 3.5. The sample includes children born between 1992 and 1997. The number of observations has decreased due to missing health outcomes-childcare attendance pairs. The dependent variables together with the set of sociodemographic controls are analytically defined in section 3.4. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

later in life. Looking closer in columns (1) and (4), the effects on healthy diet and sports go towards the expected direction (Anderson et al., 2010; Frisvold and Lumeng, 2011; Baker et al., 2015) though they remain insignificant. For instance, being eligible for entry into childcare at the age of three is associated with an increase in the probability of following a healthy diet of less than 1 percentage points if everything else is held constant (column (1), table 3.7). Comparing this coefficient to the mean value of $HealthyDiet_i$ before the reform, it suggests a negligible increase from 0.375 to 0.376 percentage points. Next, a closer look in column (3) of the same table suggests that – ceteris paribus – when the instrument $EarlierEligibility_i$ is switched on, BMI_i increases by 0.147 standard deviations. This is equivalent to a 0.035 standard deviations increase when compared to the mean value of BMI_i before the introduction of the reform. The coefficient however is not statistically significant. The same is true for the coefficients of $Smoking_i$ and $Sports_i$ (columns (2) and (4) respectively).

Table B.1.1 in the Appendix, additionally reports estimates from 2SLS regressions. The local average treatment effect is a weighted average of individuals who experience a switch in their eligibility status merely because their date of birth complies with the eligibility cutoff rule. In the absence of treatment, this group would not have been eligible. In theory, this approach would allow me to identify the causal effect of attending childcare one year earlier on long-term behavioural health outcomes if the exclusion restriction holds. However, it is debatable whether this is true in this case and thus I prefer to shift any related findings to the Appendix.

In table 3.5, I confirmed the relevance condition. In other words, $EalierEligibility_i$ is a relevant instrument for $ChildCareAttendance_i$, as being eligible for childcare entry at the age of three following the cutoff rule increases actual childcare attendance. The first stage is relevant despite the decreasing F-statistic strength once I control for child's age. Yet, the exclusion restriction denoted by the formula $Cov(EalierEligibility_i, Health_i) = 0$ might be violated. Earlier eligibility into childcare at the age of three is unlikely to be orthogonal to unobserved differences in childcare attendance and behavioural health outcomes. One reason explaining this is that the eligibility cutoff rule inducing children to enter childcare earlier might also induce children to enter elementary school earlier. Therefore, the eligibility cutoff rule might have an effect on the timing of entry into elementary school (*Grundschule*). In this case, these children would be among the youngest in their school cohort.

In an instrumental variable setting, the association between the eligibility cutoff and the elementary school cutoff threatens the validity of my two stage least squares estimates, as other factors aside from childcare may determine health status later in life. To tackle this problem, I would need to seek resort to a second exogenous shock shortly before elementary school attendance that would randomly decide the time window within which children enter school. In theory, I could isolate the childcare eligibility cutoff from the elementary school cutoff, as the former occurs around August/September and the latter in July. The main constraint is that my dataset does not have enough observations to do that. Yet, future steps could include finding a larger sample where one compares three different groups: a) those children who enter both childcare and school younger (born before July), b) those children who enter childcare younger and school later (born in August/September) and c) those children who enter both childcare and school later (born after September).

Earlier work on childcare interventions in the United States has looked into the effect on long-term health outcomes.[12] Anderson et al. (2010) compares the effect of participating

[12]Reports on the effectiveness of U.S. preschool programs on measures such as I.Q., test scores and college attendance argue for significant benefits on the treated population both in the short and in the long run (Currie, 2001; Cascio, 2009). In the long run, the positive spillovers translate into lower probability

in Head Start across siblings to show that siblings that are treated are 58% less likely to smoke as adults. Long-term effects however are more difficult to pinpoint. Frisvold and Lumeng (2011) for instance investigate the relationship between Head Start program participation and child obesity to find that although obesity is reduced in the short term, the effect vanishes after age 10.

However, this project speaks to a different from the U.S. literature (Currie, 2001; Carneiro and Heckman, 2003; Cunha et al., 2006; Fitzpatrick, 2010). Head Start was designed to improve the chances among disadvantaged kids. It aimed towards enhancing their skills and enabling them to have a more "equal footing with their more advantaged peers" (Currie, 2001). The philosophy of childcare in Europe is different. It relies on high quality *universal* access, with equal access irrespective of socioeconomic standing and where often the alternative to formal care is receiving informal home-based care. Datta Gupta and Simonsen (2015) for example use Danish survey data together with administrative records to show that being enrolled in preschool at the age of three is not associated with better cognitive outcomes upon school entry. This might be due to the quality of home care provided informally.

3.6 Discussion

In this section, I discuss alternative hypotheses that could explain my findings. First, I bring in evidence from Cornelissen et al. (2016) concerning selection into childcare and the role of family background. Then, I continue by relating my findings to the age-for-grade literature (Evans et al., 2010; Elder, 2010; Schwandt and Wuppermann, 2016).

3.6.1 The Role of Family Background

Based on work by Carneiro and Heckman (2003), among others, the main hypothesis of this paper is that being eligible for one more year of childcare attendance is associated with better long-term behavioural health outcomes (see definition in 3.4). Yet, the reduced-form estimates in table 3.7 are puzzling. An alternative explanation, emphasized by Felfe and Lalive (2015) and Cornelissen et al. (2016), hints on heterogeneous treatment effects. In this case, I would expect to see no treatment effect for children

of being a high school dropout and lower institutionalization rates. More recent U.S. Census data from three states seem to suggest that universal Pre-K availability increases the probability of enrolling to preschool by 12 to 15 percent (Fitzpatrick, 2010).

coming from a higher socioeconomic background while the effect would be positive for more disadvantaged children.

The reasons behind the heterogeneous treatment hypothesis are multi-fold. In general, children coming from better-off families may be nurtured in a higher quality home environment which includes better informal childcare. Parents, nevertheless, might choose to send their offspring to formal care in an effort to also fit their own labour market participation or leisure consumption. Havnes and Mogstad (2015) highlight that in this way parents maximize their own and their children's utility function. On the other hand, children from disadvantaged backgrounds could accrue higher benefits from the high-quality care provided in public kindergartens. Yet, the parents of those children may be less likely to enroll their offspring to childcare for two main reasons. First, mothers coming from a lower socioeconomic background are less likely to participate in the labour market. Thus, they spend more time at home. Second, parents may be less informed on the benefits of attending formal childcare in general and of the application process in specific. Especially in Germany, information on the application process, local availability of childcare and prices is not always salient. Less educated parents may thus not be as well-informed when it comes to the specifics of the application process.

In support of the aforementioned, Felfe and Lalive (2015) and Cornelissen et al. (2016) exploit the same 1996 childcare reform in Germany to look into heterogeneous treatment effects after the expansion of universal childcare. Findings point towards a reverse selection mechanism being at play, i.e. more advantaged children select into earlier childcare. Cornelissen et al. (2016) conclude that children who are profiting less from the reform are more likely to select into childcare. Felfe and Lalive (2015) also find that children of highly educated mothers are more likely to be in center-based care, mainly because of the preferential treatment they have on the waiting lists.

Furthermore, the empirical literature finds a positive relationship between expanding childcare and maternal labour supply (Blau and Currie, 2006; Havnes and Mogstad, 2011a; Black et al., 2014; Bauernschuster and Schlotter, 2015). Bauernschuster and Schlotter (2015) for instance find that childcare attendance by the youngest child may boost mother's probability of employment by 36.6 percentage points. Higher maternal labour force participation takes away time invested with the child while it contributes to the household's income. The former is found to be negatively associated with child health outcomes in the long-term (Anderson et al., 2003; Aizer, 2004; Ruhm, 2004). On the contrary, higher income is associated with better health (Deaton and Paxson, 1999). In my setting, it is difficult to argue which effect dominates or whether one effect cancels the other out.

3.6.2 Age-for-Grade Effects

Earlier entrance to childcare — and subsequently to elementary school — might be associated with adverse health effects, such as ADHD.[13] Recent empirical work has looked into the effects of being "the youngest among the schooling cohort" on health and other non-cognitive outcomes (Elder, 2010; Evans et al., 2010; Bahrs and Schumann, 2016; Schwandt and Wuppermann, 2016). Elder (2010) and Evans et al. (2010) have studied ADHD rates by exploiting school entry cutoff dates. In this paper's context, many of these children who enter childcare earlier will be the youngest in their grade. Psychologists argue that older children are more mature for learning and can accumulate skills faster. If this hypothesis holds, delaying school entry for these children could imply larger and more sustained economic returns. Schwandt and Wuppermann (2016) evaluate health insurance data from Germany to find big jumps in ADHD for birth cohorts born just before and just after the school entry cutoff date.

The German education system introduces tracking around the age of ten. The decision of the track to be followed is based on students' performance as well as the teachers' and the families' consultation. The possible alternatives include: a) high school, *Gymnasium*, b) Vocational education, *Hauptschule* or c) intermediate education, *Realschule*. The difference between the *Gymnasium* track and the other two is that the former prepares the students for university entry and hence is more prestigious. The type of educational track chosen could act as a mediator between childcare attendance and later health outcomes. My hypothesis is that earlier entry into childcare would expose children to more years of formal childcare and hence I would expect earlier entrance to be positively correlated with attending higher education, i.e. the *Gymnasium* track. In turn, children who have chosen the highest track are more likely to have better long-term health outcomes compared to children who pursue the vocational education track (Dustmann, 2004).

[13] ADHD refers to Attention Deficit Hyperactivity Disorder. It is a chronic mental condition that is mainly marked by persistent difficulty to concentrate, hyperactivity and sometimes impulsive behaviour. ADHD begins in childhood, it is particularly prevalent among boys and its effects could continue until adulthood.

In table 3.8, I look into the relationship between being eligible for earlier childcare entry (i.e. at the age of three) and the educational track followed by the individual after s/he enters school. As in section 3.5, $EarlierEligibility_i$ is a binary indicator of whether the individual was eligible for earlier childcare entry at the age of three. I run an ordered probit regression using three different education track outcomes, while controlling for a set of socioeconomic variables (see definition in section 3.4). In line with my hypothesis, a one unit increase in earlier eligibility to enter childcare (i.e. going from 0 to 1) is associated with a 0.194 increase in the log odds of attending *Higher Education* given that all other variables are held constant. The direction of the effect might be positive but it is statistically not significant (p-value=0.143, see column 1 of table 3.8). Being eligible for earlier entry into childcare is correlated with a higher chance to follow the highest education track and a significantly lower chance to follow an intermediate track (column 3). Yet, the main concern is that using the type of education track as an outcome in my regressions could be confounded with the age-for-grade effect. Many of these children who enter childcare earlier will be the youngest in their grade. For this reason, the interpretation of the findings should be done with caution.

Table 3.8: Heterogeneity — by Educational Track

Dep. var.: Education track followed	Higher Education (1)	Vocational (2)	Intermediate (3)
Earlier childcare eligibility	0.1942	-0.0717	-0.4052***
	(0.1337)	(0.1032)	(0.1284)
Federal state controls	Yes	Yes	Yes
Survey year controls	Yes	Yes	Yes
Individual level controls	Yes	Yes	Yes
Observations	1241	1241	1241
Wald Chi2	4720.42	3523.54	1913.12

Notes: The table shows ordered probit regressions using three different education track outcomes as binary dependent variables: the probability of attending higher education (column 1), the probability of following vocational training (column 2) and the probability of following intermediate education (column 3). In the German context, higher education stands for *Gymnasium*, vocational stands for *Hauptschule* and intermediate stands for *Realschule*. The explanatory variable, as before, is $EarlierEligibility_i$, a dummy taking the value of one if the individual is eligible for earlier entry into childcare. The set of sociodemographic controls included are defined in section 3.4. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

3.7 Conclusion

In this paper, I use evidence from a natural experiment in Germany — the 1996 legal claim to a place for a kindergarten — to estimate the long-term effects of being eligible for childcare entry at the age of three on behavioural health outcomes, such as smoking and exercising. This is, to the best of my knowledge, one of the first studies that use non-experimental data to examine the long-term impact of extending universal childcare on behavioural health outcomes in a European context, by exploiting a major reform. To do so, I employ individual level longitudinal data from the German Socioeconomic Panel study (SOEP). The identification strategy relies on that, after the 1996 reform was introduced, the government introduced an additional date-of-birth cutoff rule which randomly decided whether children would enter childcare exactly at the age of three (36 months old) or if they were younger than that, whether they would have to wait for one more year to entry childcare.

Using an intention-to-treat design, my hypothesis is that being eligible for earlier entry into childcare is associated with better health outcomes in the future. To explain this hypothesis, I draw reference from Carneiro and Heckman (2003) and Cunha et al. (2006) where the authors highlight the importance of early childhood interventions for long-term outcomes, among which is also health. By receiving formal care, children have the chance to get involved in various interactive activities with their peers. These may range from playing games and learning how to count and spell to interacting with teachers and more generally developing their social skills through play (Felfe and Lalive, 2015). All these factors could constitute important steps towards their emotional and cognitive development through childcare.

I document that there is no statistically significant effect of being eligible for one more year of childcare on long-term behavioural health outcomes. To explain these results, I allude to the existing literature. Evidence from Felfe and Lalive (2015) and Cornelissen et al. (2016) seem to point towards a heterogeneous treatment effect story. In particular, it could be the case that children who benefit from early childcare are coming from better-off families where the quality of informal care provided at home is already high enough and hence there is not a lot of space for improvement. In addition, the positive treatment effect of early-life investments that Carneiro and Heckman (2003) have highlighted may be cancelled out by the adverse age-for-grade effects, as argued in Evans et al. (2010), Elder (2010) and Schwandt and Wuppermann (2016).

Overall, the paper contributes to the literature about the effects of childcare attendance on long-term health outcomes. This is a highly topical question not only in economics but also in policy debates and political discussions. Supporters of early childhood interventions argue that the benefits of this type of interventions extend well beyond the children. Better-raised children is equivalent to a lower number of dysfunctional adults at public expense. Hence, it is imperative that future research establishes a causal link between longer exposure to childcare and long-tern behavioural health outcomes that could result in precise policy implications.

Appendices

APPENDIX A

Bismarck's Health Insurance and the Mortality Decline

A.1 Roll-out of Public Health Investments

Figure A.1.1: The Roll-out of Waterworks in Prussia

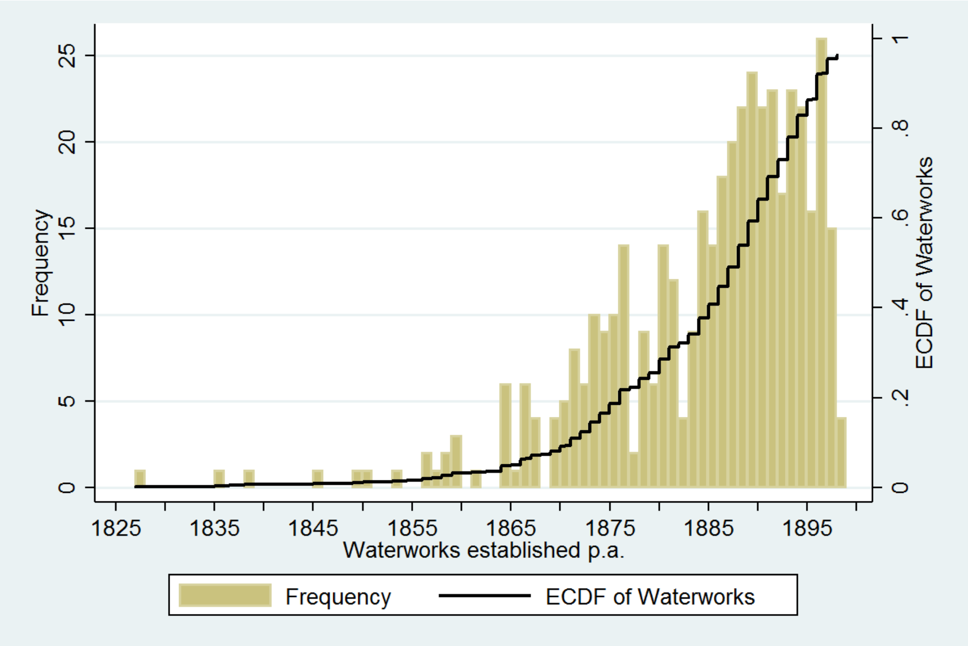
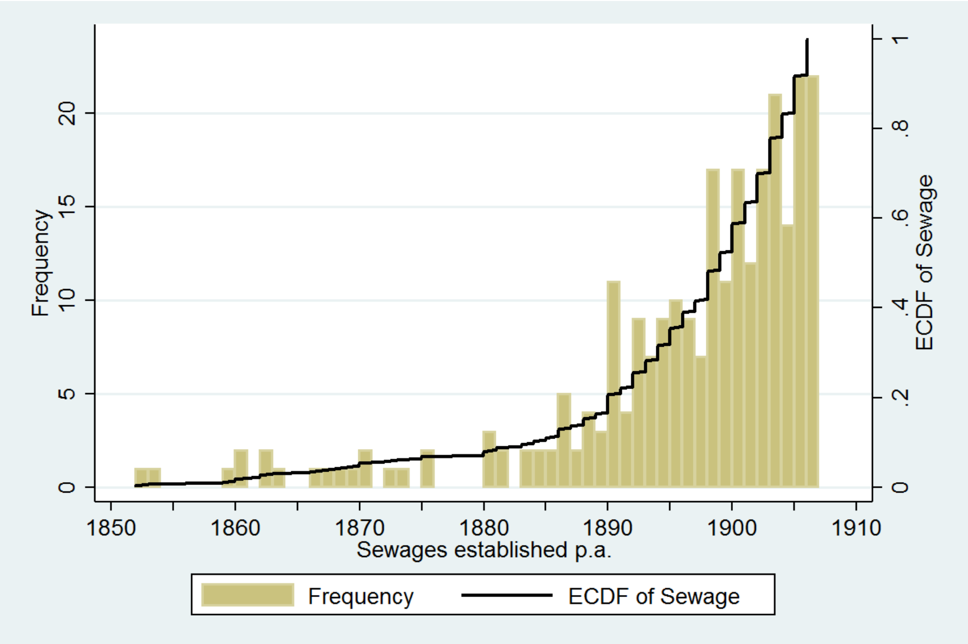


Figure A.1.2: The Roll-out of Sewerage in Prussia



A.2 Additional Robustness Checks

Figure A.2.1: Annual DiD Estimates using 1890 as Reference Year

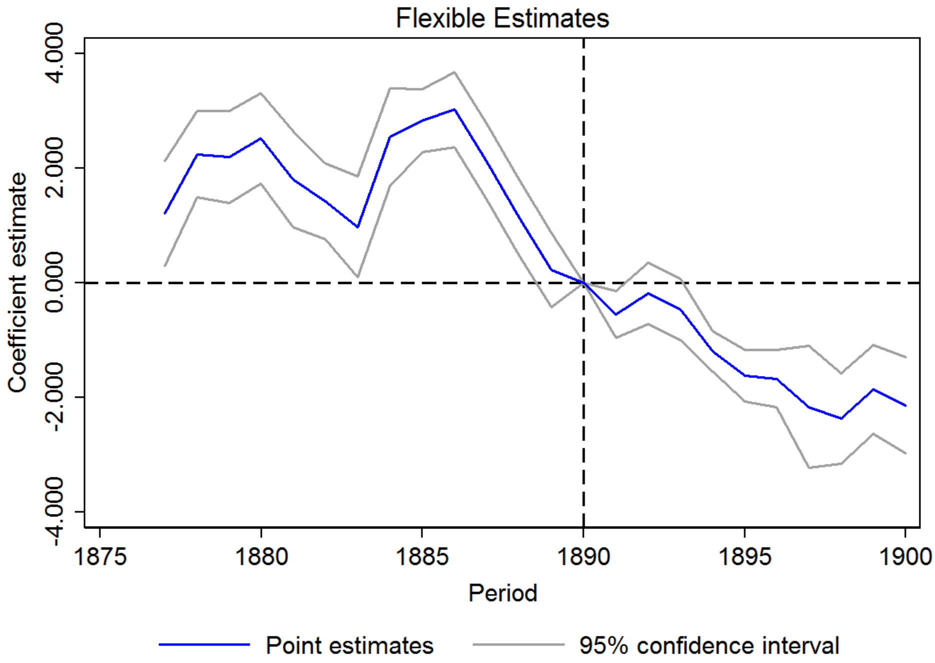
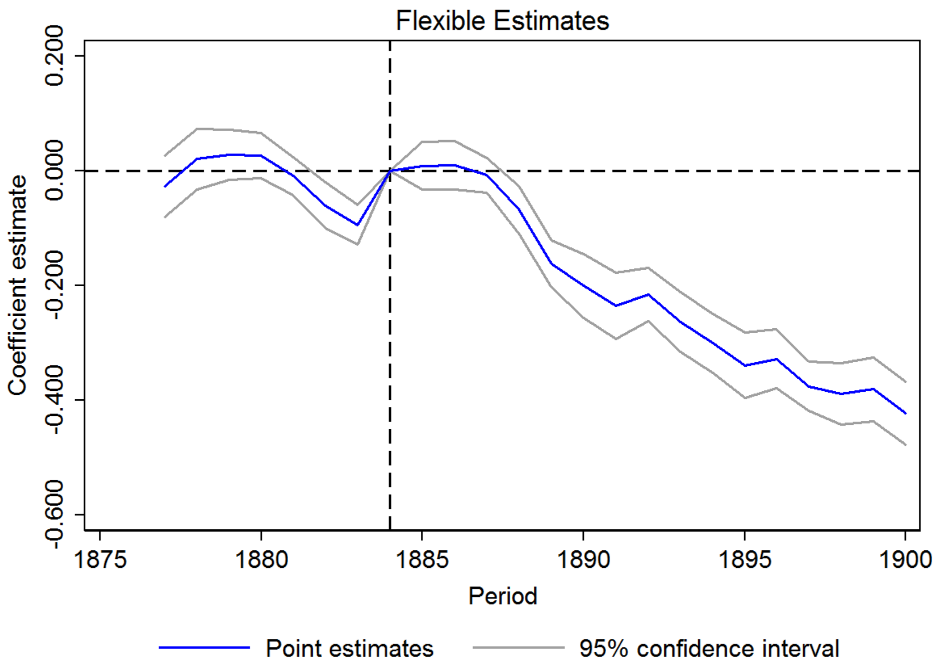


Figure A.2.2: Annual DiD Estimates in Log-specification



A.3 Health Insurance Expansion

Table A.3.1: Expansion of Health Insurance

Year	Population	Health Insured	% Insured Population
1854	17,077	254	1.5%
1860	18,136	329	1.8%
1864	19,149	458	2.4%
1868	24,069	628	2.6%
1874	25,352	715	2.8%
1885	28,232	2,263	8.0%
1890	29,819	3,457	11.6%
1895	31,697	3,998	12.6%
1900	34,254	5,123	15.0%
1905	37,058	6,192	16.7%

Notes: Population and insured in thousands. Data refer to Prussia within its respective borders. Insurance benefits vary pre- and post-1884.

Table A.3.2: Occupational Structure

	Industrial Sector			Public Sector		
	1882	1895	1907	1882	1895	1907
Total Occupational Population	9,394	12,196	16,244	1,306	1,672	2,043
Working Population	3,651	4,756	6,688	587	823	1,027
Female Working Population	585	761	1,078	61	97	166
Share in Total Population	34.4%	38.3%	43.6%	4.8%	5.2%	5.5%
Share in Working Population	31.2%	35.9%	37.1%	5.0%	6.2%	5.7%
Share in Female Working Population	17.5%	22.3%	18.7%	1.8%	2.8%	2.9%

Notes: Population in thousands. Data taken from occupation censuses conducted in the German Empire in 1882, 1895, and 1907. Sectors refer to official sector B (Industry) and E (public). *Total Occupational Population* includes children and non-employed family members and assigns to them the occupation of the father or husband, respectively.

APPENDIX B

Childcare Expansion and Behavioural Health Outcomes: Evidence from Germany

B.1 Additional Baseline Results

Table B.1.1: Reduced-form and 2SLS Results

Health outcome:	Healthy Diet		Smoking		BMI		Sports	
	RF	2SLS	RF	2SLS	RF	2SLS	RF	2SLS
Earlier childcare eligibility	0.0013		0.0693		0.1473		0.0562	
	(0.0536)		(0.0594)		(0.4548)		(0.0563)	
Childcare attendance		0.6033		0.6644		2.5160		0.2015
		(1.0153)		(0.8946)		(7.4105)		(0.9249)
Year controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Federal state controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	556	729	556	729	556	729	556	729
R squared	0.16		0.14		0.14		0.19	

Notes: The table shows reduced-form and second-stage from 2SLS estimates. As before, the dependent variables are a set of health outcomes that include: following a healthy diet, smoking, a measure of BMI and the frequency of sports. $EarlierEligibility_i$ is the main explanatory variable. The first stage is reported in table 3.5. The sample includes children born between 1992 and 1997. The number of observations has decreased due to missing health outcomes-childcare attendance pairs. The dependent variables together with the set of sociodemographic controls are defined in section 3.4. Standard errors are clustered at the individual level and are reported in parenthesis. * 10%, **5%, *** 1% confidence level. Data: SOEP, own calculations.

B.2 SOEP Data Files

B.2.1 SOEP Sample Selection

The German Socio-Economic panel dataset is a longitudinal panel survey for German households and persons that started in 1984. According to the authors of the Desktop Companion booklet Haisken-DeNew and Frick (2005), its main aim is to collect representative micro-data on households, families and individuals to track living conditions across years. A large set of variables and core questions handles topics of demography and population, health, education, labour market decisions, earnings and income as well as housing conditions. Along the years, different modules (e.g. German residents in the GDR, Immigrants, Foreigners, Innovation) are being added to the core dataset. For more details on the sample selection process, one could have a look at the institution's official process as cited by Haisken-DeNew and Frick (2005) and Kroh et al. (2015).

B.2.2 Data files used

This project has used a range of SOEP cross-section files to compile the above analysis. In detail, the personal-related Meta-dataset *PPFAD* is first combined with all subsequent datasets as it includes the unique household and individual identifier number that tracks all individuals across time. Sample characteristics specific to the survey tools employed are also included in the dataset. In my analysis, I have only accounted for private households between survey years 1987 and 2014, i.e. until the most recent SOEP data version as this paper was being written.

Information on child and maternal health along with information on parental sociodemographic variables is taken from the *P* files of SOEP's personal questionnaire. This is later combined with the *HPBRUTTO* files on gross household data. The *HPBRUTTO* files contain variables related to each household's size, the residential area of the family at the time of the survey along with detailed on the type of survey and the way the individual was contacted by the interviewer.

The *BIOAGEL* data file contains generated bio information. Among others, it includes the exact month and year of each child's birth which is crucial for my empirical analysis as my instrument is constructed based on the exact date of birth. The *BIOSIB* data file adds the number of siblings in the family on the above variables. Finally, the *PGEN* and the *KIND* files are used to supplement the analysis with detailed parental information on their educational and occupation background along with child information on the type of childcare attended. In specific, the novelty of the *KIND* data file is that it allows the researcher to know with precision if the child attended childcare at any point in his/her life. The above information is available for all survey years.

Bibliography

Acemoglu, D. and S. Johnson (2007): "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy*, 115, 925–985.

Ager, P., C. W. Hansen, and P. S. Jensen (2014): "Fertility and early-life mortality: Evidence from smallpox vaccination in Sweden," EHES Working Paper 58, University of Southern Denmark.

Aizer, A. (2004): "Home alone: supervision after school and child behaviours," *Journal of Public Economics*, 88, 1835–1848.

Alsan, M. and C. Goldin (2015): "Watersheds in Infant Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880 to 1915," *NBER Working Paper Series*.

Anderson, K. H., J. E. Foster, and D. E. Frisvold (2010): "Investing In Health: The Long-Term Impact Of Head Start On Smoking," *Economic Inquiry*, 48, 587–602.

Anderson, P., K. Butcher, and P. Levine (2003): "Maternal employment and overweight children," *Journal of Health Economics*, 22, 477–504.

Angrist, J. and A. B. Krueger (1991): "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106, 979–1014.

Ashraf, Q., A. Lester, and D. N. Weil (2008): "When Does Improving Health Raise GDP?" *NBER Macroeconomics Annual*, 23, 157–204.

Attanasio, O. P. (2015): "The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement and Policies," *Journal of the European Economic Association*, 13, 949–997.

Bahrs, M. and M. Schumann (2016): "Unlucky to Be Young? The Long-Term Effects of School Starting Age on Smoking Behaviour and Health," Working Paper. Hamburg Center for Health Economics.

Baker, M., J. Gruber, and K. Milligan (2015): "Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program," NBER Working Papers 21571, National Bureau of Economic Research, Inc.

Bartling, B., E. Fehr, and D. Schunk (2012): "Health effects on children's willingness to compete," *Experimental Economics*, 15, 58–70.

Bauernschuster, S. and M. Schlotter (2015): "Public child care and mothers' labor supply - Evidence from two quasi-experiments," *Journal of Public Economics*, 123, 1–16.

Beach, B., J. Ferrie, M. Saavedra, and W. Troesken (2016): "Typhoid Fever, Water Quality, and Human Capital Formation," *Journal of Economic History*, 76, 41–75.

Becker, G. (1975): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, National Bureau of Economic Research, Inc, second edition ed.

Belk, S. S. and W. E. Snell (1986): *The Beliefs about Women Scale:(BAWS); Scale Development and Validation*, Select Press.

Bhalotra, S., M. Karlsson, and T. Nilsson (2015): "Infant Health and Longevity: Evidence from a Historical Trial in Sweden," Discussion Paper 8969, IZA.

Black, S. E., P. J. Devereux, K. V. Løken, and K. G. Salvanes (2014): "Care or Cash? The Effect of Child Care Subsidies on Student Performance," *The Review of Economics and Statistics*, 96, 824–837.

Blau, D. and J. Currie (2006): *Pre-School, Day Care, and After-School Care: Who's Minding the Kids?*, Hanushek, E.A., Welch, F., chap. 20, 1163–1278, 2.

Bleakley, H. (2007): "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics*, 122, 73–117.

Bloom, D. E., D. Canning, and J. Sevilla (2001): "The Effect of Health on Economic Growth: Theory and Evidence," Working Paper 8587, National Bureau of Economic Research.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016): "Stereotypes," *Quarterly Journal of Economics*, 131, 1753–1794.

Bowblis, J. R. (2010): "The Decline in Infant Death Rates, 1878-1913: The Role of Early Sickness Insurance Programs," *Journal of Economic History*, 70, 221–232.

Brown, J. (1988): "Coping with the Crisis? The Diffusion of Waterworks in Late Nineteenth-Century German Towns," *Journal of Economic History*, 48, 307–318.

Bucher-Koenen, T., A. Lusardi, R. Alessie, and M. van Rooij (2014): "How financially literate are women? An overview and new insights," NBER Working Papers 20793, National Bureau of Economic Research, Inc.

Bütikofer, A., K. V. Løken, and K. G. Salvanes (2015): "Long-Term Consequences of Access to Well-child Visits," Working Papers in Economics 9/15, University of Bergen.

Card, D., C. Dobkin, and N. Maestas (2008): "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare," *American Economic Review*, 98, 2242–58.

——— (2009): "Does Medicare Save Lives?" *The Quarterly Journal of Economics*, 124, 597–636.

Carneiro, P. and R. Ginja (2014): "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start," *American Economic Journal: Economic Policy*, 6, 135–73.

Carneiro, P. and J. J. Heckman (2003): "Human Capital Policy," IZA Discussion Papers 821, Institute for the Study of Labor (IZA).

Cascio, E. U. (2009): "Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing Kindergartens into Public Schools," Working Paper 14951, National Bureau of Economic Research.

Case, A. and C. Paxson (2008): "Height, Health, and Cognitive Function at Older Ages," *American Economic Review*, 98, 463–67.

Cawley, J. and C. Meyerhoefer (2010): "The Medical Care Costs of Obesity: An Instrumental Variables Approach," Working Paper 16467, National Bureau of Economic Research.

Christelis, D., T. Jappelli, and M. Padula (2010): "Cognitive abilities and portfolio choice," *European Economic Review*, 54, 18–38.

Clark, G. and N. Cummins (2009): "Urbanization, Mortality, and Fertility in Malthusian England," *American Economic Review*, 99, 242–247.

Cleveland, W. S. (1979): "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, 74, 829–836.

Coffman, K. B. (2014): "Evidence on Self-Stereotyping and the Contribution of Ideas," *The Quarterly Journal of Economics*.

Condran, G. and E. Crimmins-Gardner (1978): "Public Health Measures and Mortality in U.S. Cities in the Late Nineteenth Century," *Human Ecology*, 6, 27–54.

Condrau, F. (2000): *Lungenheilanstalt und Patientenschicksal: Sozialgeschichte der Tuberkulose in Deutschland und England im spaeten 19. und fruehen 20. Jahrhundert.*, Vandenhoeck & Ruprecht.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schoenberg (2016): "From LATE to MTE: Alternative methods for the evaluation of policy interventions," *Labour Economics*, 41, 47 – 60, SOLE/EALE conference issue 2015.

Costa, D. (2015): "Health and the Economy in the United States, from 1750 to the Present," *Journal of Economic Literature*, 53, 503–570.

Costa, D. and R. H. Steckel (1997): *Long-Term Trends in Health, Welfare, and Economic Growth in the United States*, University of Chicago Press, 47–90.

Cunha, F., J. Heckman, and L. Lochner (2006): "Interpreting the Evidence on Life Cycle Skill Formation," in *Interpreting the Evidence on Life Cycle Skill Formation*, ed. by E. Hanushek and F. Welch, Elsevier, vol. 1, chap. 12, 697–812, 1 ed.

Currie, J. (2001): "Early Childhood Education Programs," *Journal of Economic Perspectives*, 15, 213–238.

Currie, J. and J. Gruber (1996): "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 111, 431–466.

Currie, J., M. Stabile, P. Manivong, and L. L. Roos (2010): "Child Health and Young Adult Outcomes," *Journal of Human Resources*, 45.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20, 97 – 120.

Cutler, D. M. and G. Miller (2005): "The Role of Public Health Improvements in Health Advances: The 20th Century United States," *Demography*, 42, 1–22.

Datta Gupta, N. and M. Simonsen (2015): "Academic Performance and Type of Early Childhood Care," IZA Discussion Papers 9045, Institute for the Study of Labor (IZA).

Deaton, A. (2007): "Height, Health and Development," *Proceedings of the National Academy of Sciences of the United States of America*, 104, 13232–13237.

——— (2013): *The Great Escape: Health, Wealth, and the Origins of Inequality.*, Princeton and Oxford: Princeton University Press.

Deaton, A. and C. Paxson (1999): "Mortality, Education, Income, and Inequality among American Cohorts," NBER Working Papers 7140, National Bureau of Economic Research, Inc.

DeCicca, P., D. Kenkel, and A. Mathios (2008): "Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation," *Journal of Health Economics*, 27, 904–917.

Deutsches Statistisches Bundesamt (2016): "Education, research, culture," Online, Access date: May 2016.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual Risk Attitudes: Measurement, Determinants, and Behavioural Consequences," *Journal of the European Economic Association*, 9, 522–550.

DOUGLAS, A. AND J. CURRIE (2011): "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 25, 153–72.

DOYLE, O., C. P. HARMON, J. J. HECKMAN, C. LOGUE, AND S. H. MOON (2013): "Measuring Investment in Human Capital Formation: An Experimental Analysis of Early Life Outcomes," IZA Discussion Papers 7550, Institute for the Study of Labor (IZA).

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

ECONOMIST (2016): "Election brief: Education Little Changes," Online, Access date: November 2016.

ELANGO, S., J. L. GARCÍA, J. J. HECKMAN, AND A. HOJMAN (2015): "Early Childhood Education," IZA Discussion Papers 9476, Institute for the Study of Labor (IZA).

ELDER, T. E. (2010): "The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates," *Journal of Health Economics*, 29, 641 – 656.

EUROPEAN COMMISSION (2011): "Early Childhood Education and Care: Providing all our children with the best start for the world of tomorrow," Online, Access date: June 2016.

EVANS, W. N., M. S. MORRILL, AND S. T. PARENTE (2010): "Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children," *Journal of Health Economics*, 29, 657 – 673.

EWALD, W. (1914): *Soziale Medizin. Ein Lehrbuch für Ärzte, Studierende, Medizinal- und Verwaltungsbeamte, Sozialpolitiker, Behörden und Kommunen*, Springer Berlin.

FELFE, C. AND R. LALIVE (2013): "Early Child Care and Child Development: For Whom it Works and Why," C.E.P.R. Discussion Papers 9274, Center for Economic Policy Research (C.E.R.P.).

——— (2015): "Does Early Child Care Affect Children's Development?" Working paper, St. Gallen Working Paper, revise & resubmit at Journal of Public Economics.

Felfe, C., N. Nollenberger, and N. Rodríguez-Planas (2012): "Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development," IZA Discussion Papers 7053, Institute for the Study of Labor (IZA).

Felfe, C. and L. Zierow (2015): "From Dawn till Dusk: Implications of Full-Day Care for Children's Development," Working Paper.

Fenge, R. and B. Scheubel (2014): "Pensions and Fertility: Back to the Roots– The Introduction of Bismarck's Pension Scheme and the European Fertility Decline." Working Paper 1734, European Central Bank.

Ferrie, J. P. and W. Troesken (2008): "Water and Chicago's mortality transition, 1850-1925," *Explorations in Economic History*, 45, 1–16.

Finkelstein, A. (2007): "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare," *Quarterly Journal of Economics*, 122, 1–37.

Finkelstein, A. and R. McKnight (2008): "What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending," *Journal of Public Economics*, 92, 1644 – 1668.

Fitzpatrick, M. D. (2010): "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten," *Journal of Labor Economics*, 28, 51–85.

Fogel, R. W. (2004): *The Escape from Hunger and Premature Death, 1700-2100: Europe, America, and the Third World*, Cambridge University Press.

Frevert, U. (1981): *Arbeiterexistenz im 19. Jahrhundert*, Stuttgart: Klett-Cotta, chap. Arbeiterkrankheit und Arbeiterkrankenkassen im Industrialisierungsprozess Preussens (1840-1870), 293–319.

Frisvold, D. E. and J. C. Lumeng (2011): "Expanding Exposure: Can Increasing the Daily Duration of Head Start Reduce Childhood Obesity?" *Journal of Human Resources*, 46, 373–402.

Galloway, P. R. (2007): "Galloway Prussia Database 1861 to 1914," www.patrickgalloway.com.

Geisenberger, S. (1972): "Wichtige Determinanten der Aenderung der Einkommensstruktur," in *Die Einkommensstruktur in verschiedenen deutschen Laendern 1874-1913*, ed. by H. Mueller and S. Geisenberger, Berlin: Duncker and Humblot, 77–219.

Goodman-Bacon, A. (2015): "Public Insurance and Mortality: Evidence from Medicaid Implementation," Tech. rep., mimeo.

Grahn, E. (1898-1902): *Die Städtische Wasserversorgung Im Deutschen Reiche, Sowie in Einigen Nachbarländern*, München: R. Oldenbourg.

Grossman, M. and R. Kaestner (1997): "Effects of Education on Health," *The social benefits of education*, 12, 69.

Guinnane, T. (2003): *Germany - A New Social and Economic History*, London: Hodder Arnold, vol. III, since 1800, chap. Population and the Economy in Germany, 1800-1990, 35–70.

Guinnane, T. and J. Streb (2012): "Incentives That Saved Lives: Government Regulation of Accident Insurance Associations in Germany, 1884-1914." Economic growth center discussion paper no. 1013, Yale University.

Guinnane, T. W. (2011): "The Historical Fertility Transition: A Guide for Economists," *Journal of Economic Literature*, 49, 589–614.

Guinnane, T. W. and J. Streb (2011): "Moral Hazard in a Mutual Health Insurance System: German Knappschaften, 1867-1914." *Journal of Economic History*, 71, 70 – 104.

Haisken-DeNew, J. and J. Frick (2005): "DTC: Desktop Companion to the German Socio-Economic Panel (SOEP)," Version 8.0 - Dec 2005.

Hansen, C. W., P. S. Jensen, and P. E. Madsen (2016): "Information and Disease Prevention: Tuberculosis Dispensaries," Discussion Paper 16-01, University of Copenhagen.

Hansen, C. W. and L. Lønstrup (2015): "The Rise in Life Expectancy and Economic Growth in the 20th Century," *Economic Journal*, 125, 838–852.

Hanushek, E. A. and L. Woessmann (2011): "Chapter 2 - The Economics of International Differences in Educational Achievement," Elsevier, vol. 3 of *Handbook of the Economics of Education*, 89 – 200.

Havnes, T. and M. Mogstad (2011a): "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 95, 1455–1465.

——— (2011b): "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes," *American Economic Journal: Economic Policy*, 3, 97–129.

——— (2015): "Is universal child care leveling the playing field?" *Journal of Public Economics*, 127, 100–114.

Heckman, J., R. Pinto, and P. Savelyev (2013): "Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes," *American Economic Review*, 103, 2052–86.

Hennock, E. P. (1998): "Vaccination Policy Against Smallpox, 1835-1914: A Comparison of England with Prussia and Imperial Germany," *Social History of Medicine*, 11, 49–71.

——— (2000): "The Urban Sanitary Movement in England and Germany, 1838-1914: A Comparison," *Continuity and Change*, 15, 269–296.

——— (2007): *The Origin of the Welfare State in England and Germany 1850 - 1914 - Social Policies Compared*, Cambridge University Press.

Huerkamp, C. (1985): *Der Aufstieg der Aerzte im 19.Jahrhundert vom gelehrten Stand zum professionellen Experten: Das Beispiel Preussens*, Goettingen: Vandenhoeck and Ruprecht.

Imhof, A. E. (1994[2005]): "Lebenserwartung in Deutschland, Norwegen und Schweden im 19. und 20. Jahrhundert," Za8086 data file version 1.0.0, GESIS Cologne.

Janta, B. (2013): "Caring for children in Europe," Online, Access date: August 2016.

Jappelli, T. and M. Padula (2013): "Investment in financial literacy and saving decisions," *Journal of Banking & Finance*, 37, 2779–2792.

Jayachandran, S. and A. Lleras-Muney (2009): "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines," *Quarterly Journal of Economics*, 124, 349–397.

K.A Heller, H. Kratzmeier, A. L. (1998): *Ein Handbuch zu den Standard Progressive Matrices von J.C. Raven*, vol. Bd. 1.

Kaiserlich Statistisches Amt (1884-1942): *Statistik des Deutschen Reichs, Neue Folge*, Berlin: Puttkammer & Mülbrecht.

Kesztenbaum, L. and J.-L. Rosenthal (2011): "The Health Cost of Living in a City: The Case of France at the End of the 19th Century," *Explorations in Economic History*, 48, 207–225.

——— (2016): "Sewers' diffusion and the decline of mortality: The case of Paris, 1880-1914," *Journal of Urban Economics*, forthcoming.

Khan, S. (2002): "Human Development, Health and Education," Working paper.

Kintner, H. J. (1985): "Trends and Regional Differences in Breastfeeding in Germany from 1871 to 1937," *Journal of Family History*, 10, 163–182.

——— (1987): "The Impact of Breastfeeding Patterns on Regional Differences in Infant Mortality in Germany, 1910," *European Journal of Population*, 3, 233–261.

——— (1999): "Recording the Epidemiologic Transition in Germany, 1816-1934," *Journal of the History of Medicine*, 54, 167–189.

Koch, R. (1901): "Die Bekämpfung der Tuberkulose unter Berücksichtigung der Erfahrungen, welche bei der erfolgreichen Bekämpfung anderer Infektionskrankheiten gemacht sind," *Deutsche Medizinische Wochenschrift*, 33, 549–554.

Königliches Statistisches Bureau in Berlin (1861-1934): *Preussische Statistik*, vol. 1-305, Berlin: Statistisches Bureau zu Berlin.

Kottelenberg, M. J. and S. F. Lehrer (2013): "New Evidence on the Impacts of Access to and Attending Universal Child-Care in Canada," *Canadian Public Policy*, 39, 263–286.

——— (2016): "Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Childcare," NBER Working Papers 22126, National Bureau of Economic Research, Inc.

Kroh, M., S. Kuehne, and R. Siegers (2015): "Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2014)," DIW Berlin - Deutsches Institut fuer Wirtschaftsforschung e.V.

Lavy, V. and E. Sand (2015): "On The Origins of Gender Human Capital Gaps: Short and Long Term Consequences of Teachers' Stereotypical Biases," NBER Working Papers 20909, National Bureau of Economic Research, Inc.

Lee, R., M. C. Schneider, and J. Vögele (2007): *Herausforderung Bevölkerung*, Wiesbaden: VS Verlag für Sozialwissenschaften, chap. Zur Entwicklung der Todesursachenstatistik in Preußen im 19. und frühen 20. Jahrhundert, 7–16.

Lehmann-Hasemeyer, S. and J. Streb (2016): "Does Social Security Crowd Out Private Savings? The Case of Bismarck's System of Social Insurance," Working Paper 1/2016, Priority Programme 1859: Experience and Expectation: Historical Foundations of Economic Behaviour.

Leichter, H. M. (1979): *A comparative approach to policy analysis: Health care policy in four nations*, Cambridge: Cambridge University Press.

Leonard, C. and J. Ljungberg (2010): *The Cambridge Economic History of Modern Europe*, New York: Cambridge University Press, vol. 2: 1870 to the Present, chap. Population and Living Standards, 1870-1914, 108–129.

Lleras-Muney, A. (2005): "The relationship between education and adult mortality in the United States," *The Review of Economic Studies*, 72, 189–221.

Lorentzen, P., J. McMillan, and R. Wacziarg (2008): "Death and Development." *Journal of Economic Growth*, 13, 81 – 124.

Ludwig, J. and D. L. Miller (2007): "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics*, 122, 159–208.

Lührmann, M., M. Serra-Garcia, and J. Winter (2015a): "The impact of financial education on adolescents' intertemporal choices," IFS Working Paper W14/18. Institute for Fiscal Studies.

——— (2015b): "Teaching teenagers in finance: does it work?" *Journal of Banking and Finance*, 54, 160–174.

Lusardi, A., P.-C. Michaud, and O. S. Mitchell (2016): "Optimal Financial Knowledge and Wealth Inequality," Tech. rep., Journal of Political Economy, forthcoming.

Lusardi, A. and O. S. Mitchell (2008): "Planning and Financial Literacy: How Do Women Fare?" *American Economic Review*, 98, 413–17.

——— (2014): "The Economic Importance of Financial Literacy: Theory and Evidence," *Journal of Economic Literature*, 52, 5–44.

Lusardi, A., O. S. Mitchell, and V. Curto (2010): "Financial literacy among the young: Evidence and implications for consumer policy," CFS Working Paper Series 2010/09, Center for Financial Studies (CFS).

McKeown, T. (1979): *The Role of Medicine: Dream, Mirage, or Nemesis?*, Oxford: Basil Blackwell.

Meeker, E. (1974): "The Social Rate of Return on Investment in Public Health, 1880-1910." *Journal of Economic History*, 34, 392 – 421.

Mincer, J. (1981): "Human Capital and Economic Growth," Working Paper 803, National Bureau of Economic Research.

Mokyr, J. (2000): "Why "More Work for Mother?" Knowledge and Household Behavior, 1870-1945," *Journal of Economic History*, 60, 1–41.

Murken, A. H. (1983): "Aspekte zur Begruendung der klinischen Geburtshilfe und der Entbindungskliniken im 19.Jahrundert," in *Krankenhausmedizin im 19.Jahrhundert*, ed. by H. Schadewaldt and J. H. Wolf, Munich: Muenchener Vereinigung fuer Geschichte der Medizin, 70–97.

Niederle, M. and L. Vesterlund (2010): "Explaining the Gender Gap in Math Test Scores: The Role of Competition," *Journal of Economic Perspectives*, 24, 129–44.

Ogasawara, K. and G. Kobayashi (2015): “The Impact of Social Workers on Infant Mortality in Inter-war Tokyo: Bayesian Dynamic Panel Quantile Regression with Endogenous Variables,” *Cliometrica*, 9, 97–130.

Organisation for Economic Co-operation and Development (2014): “Enrolment in childcare and pre-schools,” Online, Access date: September 2016.

Preston, S. H. (1975): “The Changing Relation between Mortality and Level of Economic Development,” *Population Studies*, 29, 231–248.

Robins R. W., Hendin, H. M. and K. H. Trzesniewski (2001): “Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale,” *Personality and Social Psychology Bulletin*, 27, 151–161.

Robinson, P. M. (1988): “Root-N-Consistent Semiparametric Regression,” *Econometrica*, 56, 931–954.

Rosenberg, H. (1967): *Depression und Bismarckzeit: Wirtschaftsablauf, Gesellschaft und Politik in Mitteleuropa*, Veroeffentlichungen der Historischen Kommission zu Berlin.

Ruhm, C. J. (2004): “Parental employment and child cognitive development,” *Journal of Human Resources*, 39, 155–192.

Ruhm, C. J. and J. Waldfogel (2012): “Long-Term Effects of Early Childhood Care and Education,” *Nordic Economic Policy Review*, 23–51.

Salomon, H. (1906-1907): *Die Städtische Abwässerbeseitigung in Deutschland*, vol. 2, Jena: G. Fischer.

Schwandt, H. and A. Wuppermann (2016): “The youngest get the pill: ADHD misdiagnosis in Germany, its regional correlates and international comparison,” *Labour Economics*.

Sieveking, G. H. (1903): *Die Milch und Ihre Bedeutung fuer Volkswirtschaft und Volksgesundheit*, Von C. Boysen Hamburg.

Spree, R. (1996): *Einem jeden Kranken in einem Hospitale sein eigenes Bett: Zur Sozialgeschichte des Allgemeinen Krankenhauses in Deutschland im 19.Jahrhundert*, Frankfurt: Campus, chap. Quantitative Aspekte der Entwicklung des Krankenhauswesens im 19. und 20.Jahrhundert: Ein Bild innerer und aeusserer Verhaeltnisse, 51–90.

Strittmatter, A. and U. Sunde (2013): "Health and Economic Development–Evidence from the Introduction of Public Health Care." *Journal of Population Economics*, 26, 1549 – 1584.

Subramanian, S. and A. Deaton (1996): "The Demand for Food and Calories," *Journal of Political Economy*, 104, 133–162.

Szreter, S. (1988): "The Importance of Social Intervention in Britain's Mortality Decline C.1850-1914: A Re-interpretation of the Role of Public Health," *Social History of Medicine*, 1, 1–38.

Tennstedt, F. (1983): *Vom Proleten Zum Indusriearbeiter: Abeiterbewegung und Sozialpolitik in Deutschland 1900 bis 1914*, Otto Brenner Stiftung.

Thomasson, M. (2013): "Economic History and the Healthcare Industry," in *Routledge Handbook of Modern Economic History*, ed. by R. Whaples and R. E. Parker, New York: Routledge, chap. 16, 177–188.

van den Berg, G. J. and B. Siflinger (2016): "The Effect of a Universal Child Care Reform on Child Health- Evidence from Sweden," Working Paper.

van Rooij, M., A. Lusardi, and R. Alessie (2011): "Financial literacy and stock market participation," *Journal of Financial Economics*, 101, 449 – 472.

Vögele, J. (1998): *Urban Mortality Change in England and Germany, 1870-1913.*, Liverpool University Press.

Weil, D. N. (2007): "Accounting for the Effect of Health on Economic Growth." *Quarterly Journal of Economics*, 122, 1265 – 1306.

Winegarden, C. and J. E. Murray (1998): "The Contributions of Early Health-Insurance Programs to Mortality Declines in Pre-World War I Europe: Evidence from Fixed-Effects Models," *Explorations in Economic History*, 35, 431–446.

World Bank (2016a): "SABER: System Approach for Better Education Results," Online, Access date: August 2016.

——— (2016b): "The World Bank Database," Online, Access date: August 2016.

World Health Organization (2016a): "Health effects of smoking among young people," Online, Access date: September 2016.

——— (2016b): "Vaccination greatly reduces disease, disability, death and inequity worldwide," Online, Access date: October 2016.

Wüst, M. (2012): "Early Interventions and Infant Health: Evidence from the Danish Home Visiting Program," *Labour Economics*, 19, 484–495.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht.
Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

Datum: 06.12.2016

Unterschrift: Anastasia Driva