Health Inequalities at Different Stages of the Lifecycle

Sophie-Charlotte Meyer

Wuppertal 2016

Health Inequalities at Different Stages of the Lifecycle

Sophie-Charlotte Meyer

Inauguraldissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaft (Dr. rer. oec.)
an der Fakultät für Wirtschaftswissenschaft
— Schumpeter School of Business and Economics —
der Bergischen Universität Wuppertal

vorgelegt von
Sophie-Charlotte Meyer
aus Menden (Sauerland)

Wuppertal, den 8. Juni 2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170118-123308-9
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20170118-123308-9

Erstgutachter: Prof. Dr. Hendrik Jürges

Zweitgutachterin: Prof. Dr. Kerstin Schneider

Tag der mündlichen Prüfung: 05. Dezember 2016

Für meinen Vater.

Danksagung

Die vorliegende Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Gesundheitsökonomie und -management der Bergischen Universität Wuppertal und wäre ohne die Unterstützung vieler Personen, denen ich im Folgenden danken möchte, nicht möglich gewesen.

Zunächst möchte ich mich sehr herzlich bei meinem Doktorvater und Koautor, *Prof. Hendrik Jürges* bedanken. Während meiner gesamten Promotion hat er sich immer Zeit genommen ausführlich mit mir über meine Anliegen zu diskutieren. Seine kritischen Anmerkungen, die vielen wertvollen Anregungen sowie seine hervorragende und umfassende Unterstützung waren für den Erfolg der Dissertation und nicht zuletzt auch für die erfolgreiche Publikation des zweiten Kapitels essentiell.

Für die Übernahme des Zweitgutachtens und den konstruktiven Kommentaren insbesondere während der Disputation danke ich *Prof. Kerstin Schneider*. Daneben möchte ich mich vielmals bei *Prof. Juliane Köberlein-Neu*, *Prof. Jörg-Peter Schräpler* sowie *Prof. Ina Garnefeld* für die kritische und anregende Diskussion während meiner Disputation bedanken.

Meinen derzeitigen und ehemaligen Kolleginnen und Kollegen des Lehrstuhls sowie des Bergischen Kompetenzzentrums für Gesundheitsökonomik und Versorgungsforschung danke ich für die freundschaftliche Arbeitsatmosphäre und stete Hilfsbereitschaft. Besonders möchte ich mich bei meinen Weggefährten *Lars Thiel* und *Marianna Schaubert* sowie *Tetyana Bruditz* für ihre wertvolle Hilfe und Unterstützung bedanken, die wesentlich zum Gelingen dieser Arbeit beigetragen haben. Meiner Koautorin *Annemarie Künn-Nelen* danke ich in diesem Zusammenhang für die gute Zusammenarbeit und Unterstützung insbesondere zu Beginn meiner Promotion.

Danken möchte ich auch den vielen Personen, die ich während verschiedener Konferenzen und Workshops kennenlernen durfte. Ihre Anmerkungen und Diskussionen haben die Qualität meiner Dissertation maßgeblich beeinflusst. Namentlich möchte ich mich bei *Jan Marcus* und *Francesca Carta* für die konstruktiven Anregungen im Rahmen ihrer Koreferate bedanken. Den Mitarbeitern beim Forschungsdatenzentrum des IT.NRW danke ich für die Bereitstellung der Daten. Mein besonderer Dank gilt hierbei *Frau Dr. Kerstin Schmidtke* für ihr Interesse an meinen Untersuchungen.

Meinen Freunden danke ich vom ganzen Herzen für ihr liebevolles Verständnis und die aufmunternde Ablenkung, ohne die ich einen solchen Arbeitsumfang nicht hätte bewältigen können. *Manuel Pasvoß* danke ich besonders für seine Unterstützung und die mehrfache Durchsicht meiner Arbeit.

Mein ganz besonderer Dank aber gilt meiner Mutter *Karin Meyer* sowie meiner Schwester *Nadine Spöth* und ihrer Familie für ihr Vertrauen in mich und ihre bedingungslose Unterstützung, ohne die diese Arbeit nicht möglich gewesen wäre.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1 Motivation

Socioeconomic inequalities in health and health behavior seem to be omnipresent: Regardless of country and time, individuals with higher socioeconomic status experience on average a healthier and longer life compared to those with lower socioeconomic status (e.g. Deaton, 2003; Smith, 2004; Mackenbach et al., 2008; Mirowsky and Ross, 2003; Cutler and Lleras-Muney, 2006). Among the different indicators of socioeconomic status, which are typically the level of education, occupation, and income, education is outstanding not least because it has a determining influence on the other aspects. The existence of socioeconomic disparities in health, also referred to as health inequalities, is well-documented across various research disciplines, such as epidemiology, sociology, economics or demography. By contrast, it is less clear why this relationship persists and empirical evidence on mechanisms is yet largely inconsistent. In general, the literature offers three explanations. First, socioeconomic status might have a causal effect on health. For instance, education might influence the realization and processing of medical instructions or health-related information (Grossman, 2006). Second, there might be reverse causality, i.e. a causal link running from health to socioeconomic status. For instance, adverse health (shocks) might impair the educational attainment or labor market outcomes, such as income or employment (Currie, 2009; Cutler and Lleras-Muney, 2010). Third, both health and socioeconomic status might be determined by third factor variables, such as (time) preferences or family resources, that rather imply selection than causation. As (reverse) causation and selection are likely concurrent within these complex links, it is challenging to disentangle socioeconomic disparities in health. In his recent review, Grossman concludes: “There is enough conflicting evidence in the studies that I have reviewed to warrant more research on the question of whether more schooling does in fact cause better health outcomes” (2015, p.14). Recent evidence suggests that selection largely accounts for the relationship between socioeconomic status and health or health behavior (Conti and Heckman, 2010; Conti et al., 2010; Von Hippel and Lynch, 2014; Maralani, 2013).

Reducing socioeconomic inequalities in health is of particular importance for welfare states that aim to provide equal opportunities for these important aspects of life. However, knowing about the (causal) direction and mechanisms is crucial to implement effective programs in different policy areas, such as health, social or educational policy. For instance, if schooling induces individuals to be healthier, expanding the access to higher education for individuals with poor socioeconomic resources would be successful in reducing these disparities. In contrast, these actions would likely be ineffective, if the opposite direction was true and (child) health affected educational attainment. In this case, policy actions aiming to improve child health and lifestyles would be more likely to successfully reduce educational differences in health.[1] The inconsistent evidence indicates that there is no such uniform relationship between these two life parameters. Rather, these complex links interacting with childhood conditions, preferences and other (contextual) factors seem to vary over the health outcome studied and the individual's life course. The literature largely focused on health inequalities and its mechanisms in adulthood for a long time. More recently, researchers have focused on inequalities at different stages of the life course which might be a promising way to unravel these links. Although health problems increasingly appear when individuals get older, the causes might trace back to earlier periods of life (Barker, 1998; Blane, 2006). Socioeconomic disparities often persist or even cumulate over the individual's life course. Assessing when socioeconomic disparities in certain health conditions or health behavior manifest is thus important to better detect the specific mechanisms acting at different stages of life (Smith, 2004). To illustrate, let us take the educational differences in smoking as an example. Individuals usually take up smoking in adolescence with the lower educated starting more frequently than the higher educated. Smoking is highly relevant for later health as it is the leading behavioral cause of death in Germany and other industrialized countries. An effective policy action thus aims to reduce the smoking initiation rates in this period of life. Knowing about the timing when specific inequalities in health (behavior) manifest might take us a step forward to implement effective political programs in different areas of life, such as school or the workplace, to provide equal health opportunities for all subpopulations. This thesis draws on this previous literature by focusing on specific aspects of socioeconomic disparities in health at different stages of the life course.

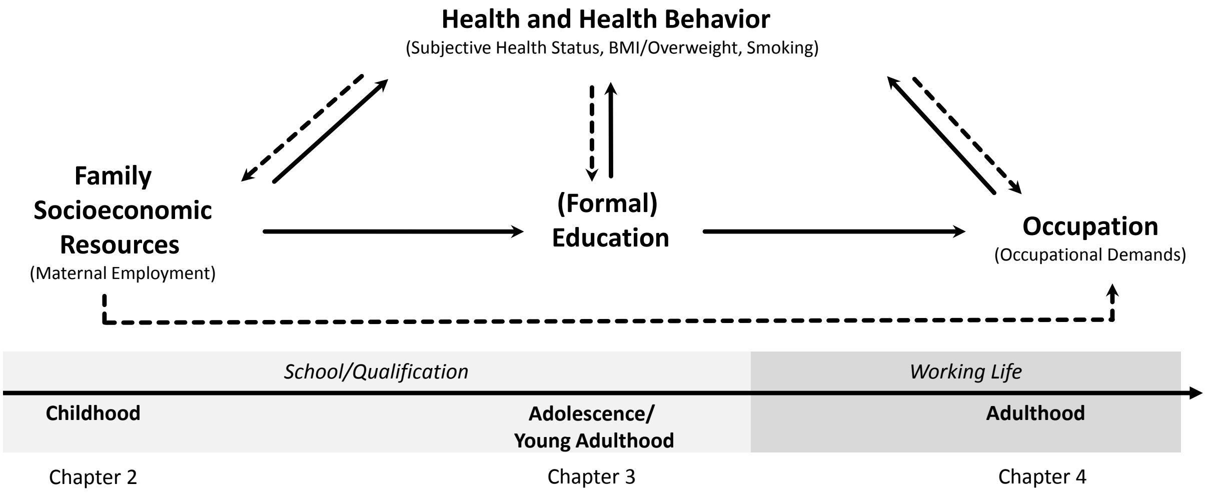
1.2 Aim and Contribution

The main purpose of this thesis is to contribute to our understanding of these complex links between socioeconomic status and health (behavior). Specifically, this thesis focuses on health inequalities and its mechanisms at different stages of the life cycle. However, given the complex links it is impossible to explore health inequalities across the whole life course in a comprehensive way. For that reason, this dissertation aims to shed light on three specific aspects of these health inequalities. Figure 1.1 conceptualizes the theoretical framework connecting these specific aspects.

Despite genetic disposition, family socioeconomic resources, such as parental education, employment or income largely determine a child's health and health behavior. For instance, maternal employment as one specific aspect of socioeconomic status, likely increases family income but also reduces the available time for the family. Both might be

[1]The fact, that we still know too little about the underlying mechanisms might also explain the modest success of the policy actions that have been taken to reduce health inequalities in the UK (see Mackenbach, 2011).

Figure 1.1: Conceptional framework of the thesis



Note: Dashed arrows present relationships that are not directly investigated in this thesis.

related to child health outcomes, such as overweight. In Chapter 2, this thesis relates to this trade-off and tries to answer the following questions: Does maternal employment promote childhood overweight and if so, why?

Parental education and other family resources do not only determine child health (behavior) but also the child's own educational attainment. The formal educational degree is usually obtained in (late) adolescence. During that time, harmful behaviors, such as cigarette smoking, are shaped as well. The formal educational degree, which is most often measured in surveys, is thus likely obtained after individuals take up smoking. This renders a strong causal effect of formal education, as it is often stated in the literature, unlikely. Chapter 3 draws on this observation and approaches the following questions: Can the relationship between formal education and smoking be interpreted as causal? What is the role of mechanisms operating during school? Is a specific health education rather than formal education related to smoking decisions?

The educational attainment determines the individual's professional career due to formal regulations but also individual preferences. The occupational choice and the related demands within an occupation might in turn also be related to the individual's health and health behavior, such as BMI and smoking. Occupations and more specifically, occupational demands might thus (partly) mediate educational disparities in health and health behavior in adulthood. Chapter 4 is related to this observation and tries to answer the following questions: Do occupational demands mediate the educational gradient in health (behavior) and if so, to what extent?

Each chapter adds to the interdisciplinary literature on socioeconomic disparities in health (behavior) in its own way. The specific contribution of each study is presented in the following chapters in more detail. However, there are two common contributions across the three studies. First, all studies go beyond estimating the mere relationship between the analyzed socioeconomic determinant and health (behavior) putting a particular focus on detecting mechanisms. Starting from a pure correlation, causation is often implicitly assumed although there is lacking evidence on the underlying mechanisms. But knowing about causality and the responsible mechanisms is crucial in order to implement policy measures to effectively reduce socioeconomic inequalities. During the last decades identifying causal effects has come into sharp focus within the empirical economics literature. For that reason, a special attention in this thesis is given to causality issues, although only Chapter 2 applies an advanced econometric technique (IV estimation) to empirically identify a causal effect. However, all of the three chapters provide a detailed discussion on the assumptions and potential biases which might occur due to unobserved heterogeneity or reversed causality.

The second contribution of this dissertation relates to the data sources used. Throughout all chapters the main analyses are based on the German Microcensus, an annual survey representative for the German population which is collected by the German Federal Statistics Office. The Microcensus is the largest annually-conducted household survey in Germany and covers 1 % of the German households. Although it has rarely been used to study social disparities in health in the past, it has several advantages over other available data sources. The data include diverse information on the individuals' socioeconomic conditions, such as employment or education, which are asked every year. Supplementary questions on health are included every four years and contain information on anthropometric measures, (retrospective information on) smoking behavior, and health status. Although the German Microcensus is cross-sectional, it has a huge potential due to its large sample size.[2] While the answers to the health questions are voluntary, participation in the Microcensus is basically mandatory leading to a relatively low non-response rate also for the health-related questions. The large number of observations not only increase the precision of estimates but also enables us to perform differentiated analyses to detect socioeconomic patterns and cohort trends. For instance, in Chapter 3 we estimate educational differences in smoking by year-of-birth. Differentiated measures of education and occupations are exploited in Chapter 4. The large sample also enables us to perform robustness analyses for specific subpopulations that are important to explore the validity of our results. Such analyses are hardly possible with other German data sources, such as the German Socio-Economic Panel, as differentiated and stratified analyses often lack statistical power. Moreover, the German Microcensus is conducted as a household survey and thus provides information on the household and the individual but also on families. In Chapter 2 we base our analyses on the family level and relate family charac-

[2]The data are collected through a rotating procedure so that individuals in the same household are interviewed up to four times. However, because the health-related questions are asked in a four-year cycle (except for the years 2003 and 2005), we are not able to identify individuals across waves for the purpose of our analyses.

teristics to child overweight. Apart from the limited health-related measures within the Microcensus, the data also include little detailed information on specific spheres, such as individual lifestyle behavior or occupational demands. In order to complement the analyses, we additionally draw on other representative data sources that include detailed information on the specific facet explored. In Chapter 2 and 3 we use the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) to explore potential mechanisms of socioeconmic disparities in smoking and overweight operating during childhood and adolescence. In Chapter 4, we combine the German Microcensus data with the BIBB/BAuA-Employment Survey which includes detailed information on the individual's work environment and occupational demands. The combination of the datasets allows us to explore the importance of occupational demands for educational differences in health (behavior) among the working population. Each of these data sources has distinct advantages for the research question at hand that complement our analyses.

1.3 Outline

The thesis contains three self-contained chapters which deal with socioeconomic disparities in health and health behavior at different stages of the individual's life cycle. The following subsections briefly summarize the three chapters and discuss the main findings.

Chapter 2: Maternal Employment and Childhood Overweight

The second chapter (published in Economics and Human Biology) is single-authored and addresses maternal employment as one possible determinant of childhood overweight. The empirical evidence regarding the direction of the relationship between maternal employment and child overweight is inconsistent. A widespread finding among studies from the US and the UK is that maternal employment is correlated with an increased risk of child overweight, even in a causal manner. In contrast, studies from other countries obtain less conclusive results. As evidence for Germany is still scarce, the purpose of this chapter is to identify the effect of maternal employment on childhood overweight in Germany and explore potential underlying mechanisms that might explain the relation between maternal employment and child overweight. The analyses are restricted to children aged 9–12 years, a crucial age range which has yet rarely been studied in the context of maternal employment. In order to address the potential endogeneity, an instrumental variable approach is applied using the number of younger children in the household as instrument for maternal employment. The number of younger siblings in the household is likely to be a good instrument for maternal employment when controlling for a detailed range of maternal and family characteristics, as it is strongly negatively related to maternal employment. Besides the investigation of the mere relationship between maternal employment and child weight, the IV approach is used to explore the role of dietary and activity habits as mechanisms.

OLS models state that maternal full-time employment is related to a 5 percentage point higher probability of the child to be overweight. In contrast, IV estimates suggest that maternal full-time employment increases the probability of the child to be overweight by 30 percentage points. This suggests that the simple correlation between maternal employment and child overweight likely understates the effect as mothers may select themselves into full-time employment due to unobserved characteristics (e.g. maternal effort, ability) that promote child health. These results are very similar across the two data sources. Based on the KiGGS data, the results further indicate that maternal full-time employment promotes unhealthy dietary habits, namely a lower consumption of fruits and vegetables, and a higher consumption of soda drinks and processed food. Children of full-time employed mothers also spend more time watching TV and playing video games. These dietary and sedentary habits might be relevant mechanisms through which maternal employment promotes child overweight. Based on the German Microcensus, descriptive OLS results suggest that mothers' nonstandard work schedules, such as working on Saturdays, Sundays or at nights, are also related to child overweight. Although there are limitations to our IV approach, several sensitivity analyses confirm the robustness of our findings.

Taken as a whole, the implications of the findings of Chapter 2 are not that mothers should drop out of the labor force nor that maternal employment is per se harmful for child health. Public policy should rather invest in ways improve the job conditions of working parents and facilitate access to or provide healthy food and activities for children. This is especially relevant as daycare arrangements for secondary-school-aged children are still limited in Germany.

Chapter 3: Revisiting the Relation between Education and Smoking

The third chapter is joint work with Hendrik Jürges and addresses educational differences in smoking during adolescence and young adulthood. Starting from the well-documented correlation between education and smoking, researchers have recently tried to isolate its causal effect by exploiting different sources of exogenous variation in education such as changes in compulsory schooling – with mixed results. Some studies find a protecting effect, others find no evidence that education has a causal effect on smoking behavior. This may partly be explained by the fact that the estimated effects are confined to different (often small) subpopulations affected by a policy change. Moreover, empirical evidence on possible mechanisms (e.g. health knowledge) is scarce. The aim of this chapter is to descriptively investigate whether post-compulsory education can have a causal effect on smoking for a broader population. Moreover, we aim to explore whether health education, one of the key explanations discussed in the literature, is the predominant pathway and what mechanisms operating during school might be relevant.

We argue that if the relationship was causal, educational differences would be absent at smoking initiation which typically occurs in early adolescence and thus before post-compulsory schooling is completed. Making use of retrospective information on the age at smoking onset included in the German Microcensus data, we employ discrete time event history models to compare age-specific hazard rates for individuals with compulsory and post-compulsory education. To assess cohort changes, we perform all analyses separately by birth cohort and sex. Based on another data source (KiGGS), we describe the role of certain mechanisms operating prior to or during school, such as family resources (e.g. parental smoking and socioeconomic background) or potentially endogenous characteristics (e.g. peer smoking and subjective well-being). Further, we explore the importance of health knowledge indicated by the individual's medical studies or health-related training. Specifically, we assess whether doctors and nurses are more likely to stop smoking (causal effect) and whether these individuals were already less likely to take up smoking before learning their occupation (selection), compared to individuals with the same educational level.

The data show that educational differences are already apparent at smoking initiation. About 85 % (93 %) of the educational differences in smoking among men (women) are determined before the age of 16. Whether an individual ever smokes is thus predominantly determined at an age before compulsory schooling is completed. Especially smoking peers seem to be relevant for the decision to take up smoking. Health education (knowledge) itself unlikely has a (strong) causal effect on smoking decisions.

In conclusion, the results are incompatible with the widespread finding that formal education has a strong causal effect on smoking behavior. Rather, it is more likely that (unobserved) factors determining both the selection into smoking and education and resulting peer effects, are responsible for educational differences in smoking.

Chapter 4: Do Occupational Demands Mediate the Educational Gradient in Health (Behavior)?

The fourth chapter is co-authored by Annemarie Künn-Nelen and analyzes the relationship between education and health (behavior) during the working life. We focus on occupational demands as specific mechanism in the education-health (behavior) relationship. The extent to which workers are distributed across occupations, and even the extent to which they are exposed to different occupational demands, clearly depends on their education level. As different physical and psychosocial occupational demands are expected to affect health as well, occupational demands might serve as a mechanism.

Merging the German Microcensus 2009 data with the BIBB/BAuA-Employment Survey 2005/2006, we investigate to what extent occupation-specific demands mediate educational differences in subjective health and health behavior (BMI and smoking). We consider ergonomic, environmental, psychological, social and time occupational demands. Based on the BIBB/BAuA-Employment Survey 2005/2006, we estimate these occupational demands at the occupational level taking differences across sex and age groups into account. The advantage of this method is that we are able to get rid of individual characteristics such as personality traits which affect both reported health and perceived occupational demands. Finally, these demands are aggregated and merged via occupations, age and sex to the Microcensus data. The mediation analyses are based on the entire working population aged 25–65 years. Omitting the group of currently not employed individuals (as it is usually done) likely biases the results. That is, because these individuals might have quit the labor force due to occupational demands with adverse health effects.

We find that occupational demands are significantly related to subjective health and health behavior. Whereas ergonomic, environmental and social demands are positively correlated with the health outcomes, psychological demands are negatively related and time demands appear to be unrelated to health (behavior). In a regression of health on education including the occupational demands, our results indicate that occupational demands mediate educational differences in subjective health status for lower educational levels only. Regarding the health behavior considered, this partial mediation is more comprehensive. Education coefficients on BMI and smoking significantly reduce up to 21 % and 27 % when the occupational demands are included. Especially social demands seem to be crucial for the relationship between education and health behavior.

Overall, our findings indicate that existing inequalities in working conditions do matter for the educational gradient, especially in BMI and smoking. Improving the working conditions especially for lower educated individuals might thus contribute to reduce educational differences in health. Moreover, this study provides important suggestive evidence that there might be dynamic effects in explaining the relation between education and health (behavior) via work-related conditions, such as occupational demands.

CHAPTER 2

Maternal Employment and Childhood Overweight

Published in *Economics and Human Biology* as: Meyer, Sophie-Charlotte (2016). Maternal Employment and Childhood Overweight in Germany. *Economics and Human Biology* 23: 84-102, DOI: 10.1016/j.ehb.2016.05.003.

2.1 Introduction

Childhood overweight has recently become a major public health concern, not at least because overweight children are at an increased risk to become overweight adults which is in turn associated with a number of health and social problems (see Cawley 2010, 2011 for overviews). Given the rising share of overweight children in nearly all developed countries, research has focused on various determinants of child overweight, apart from the major direct causes, excessive caloric consumption and lack of physical activity. As also female (and maternal) labor supply has been rising during the last decades, recent studies explore whether maternal employment is a determinant of child overweight. Theoretically, maternal employment implies a trade-off between time and money. On the one hand, employment reduces the mother's time to care for her children which might promote an unhealthy lifestyle with reduced physical activity or low-quality/processed food. On the other hand, maternal employment increases family income which might enable parents to invest in high-quality child care or healthy food. Consequently, the direction of the relationship between maternal employment and child overweight is a priori unclear and a positive as well as a negative relationship is conceivable. The empirical evidence regarding the direction of the relationship between maternal employment and child overweight is inconsistent. Moreover, most of the existing studies focus on maternal employment across a wide age range or the entire childhood (e.g. Fertig et al., 2009; Nie and Sousa-Poza, 2014). A widespread finding among studies from the US and the UK is that maternal employment is correlated with an increased risk of child overweight, even in a causal manner (e.g. Anderson et al., 2003; Ruhm, 2008 for the US; von Hinke Kessler Scholder, 2008 for the UK). These findings are in contrast to studies from other countries which obtain less conclusive results. Whereas a negative effect has been found for Denmark using an instrumental variable (IV) strategy (Greve, 2011), no relationship appears in cross-sectional data from China (Nie and Sousa-Poza, 2014) and selected European regions (Gwozdz et al., 2013).

Evidence on the relationship between maternal employment and child weight for Germany is scarce, although Germany can be regarded as a particularly interesting case for several reasons. Traditionally, the male-breadwinner model has been encouraged by German policy, e.g. by a tax law that favors families with one high earner. Nevertheless, the share of employed mothers has been rising in Germany as well, by 7 percentage points up to 66 % between 2000 and 2012 (BMFSFJ, 2012).[3] However, the reconciliation of family and work has long been neglected.

In this chapter we aim to determine the effect of maternal employment on child overweight in Germany using the number of younger siblings in the household as an instrument. Further, we explore potential underlying mechanisms that might explain

[3] Female labor participation rates are traditionally higher in East-Germany.

this relationship. Exploring mechanisms is of particular importance as it reveals potential factors which might contribute to the prevention of childhood overweight. We replicate previous findings for Germany and add to the existing literature in several ways. First, we focus on children in preadolescence (9 to 12 years old), which is found to be an important development stage.[4] Children of this age are more likely to co-determine their allocation of time compared to younger children who depend more on their parents' decisions. Another reason for the focus on this specific age group is that in Germany, children typically switch schools from primary to secondary school at about 10 years of age. This transition involves a range of changes for the children. Maternal employment might affect the child's well-being especially during that stage of life. However, evidence on the effects of maternal employment in preadolescence is still scarce and mostly confined to the US (e.g. Morrissey et al., 2011; Ruhm, 2008).

Second, we aim to determine the effect on child overweight taking the endogeneity of maternal employment into account. In order to deal with the selection into maternal employment, we use an IV approach using the number of younger siblings in the household.[5] A related instrument, the youngest sibling's kindergarten eligibility age, has been previously used to instrument maternal employment (Morrill, 2011). In this study, we argue that the number of younger siblings in the household is likely to be a good instrument when controlling for a range of child and family characteristics, because it is strongly negatively related to maternal employment. With respect to the validity of the instrument chosen, we provide a theoretical discussion on the instrument and perform several robustness analyses. We built our analyses on two data sources, each having distinct advantages. Using several cross-sections of the German Microcensus enables us to estimate the relationship for a large number of children and thus obtain estimates of higher precision. This is also crucial for several robustness analyses aiming to explore the validity of the instrument introduced in this study. However, the health-related information in the German Microcensus is sparse. We thus base or analyses on a second data source, the German Health Interview and Examination Survey for Children and Adolescents (KiGGS).[6] Although the KiGGS survey is also cross-sectional to date, it has considerable advantages over the other data source. Body height and weight are measured by trained medical staff and the data include information on birth weight and circumstances during pregnancy. This variables are helpful in order to mitigate any potential bias due to mothers' employment responses from child health at birth.

[4]However, the relationship can be found for younger and older children as well (see Table 2.A.8 in the Appendix).

[5]Garcia et al. (2006) use the same instrument among other family characteristics when estimating the effect of maternal employment on child overweight using Spanish data. However, they simply claim that the exogeneity of maternal employment cannot be rejected. The paper neither provides any results of the first stage or second stage, nor does it include a discussion on the instruments.

[6]The data has already been used to examine social determinants of child health (e.g. Reinhold and Jürges, 2012).

Our final contribution lies in the exploration of possible underlying pathways. We focus on mechanisms that have already been investigated but led to mixed results. The KiGGS data are unique as they include detailed information on the child's lifestyle behavior and conditions which could potentially explain why maternal employment promotes childhood overweight. Specifically, we consider unhealthy dietary habits, lack of physical activity and media consumption.[7] Applying our IV strategy we explore whether these habits might be induced by maternal employment to some extent. Thereby, we aim to yield further insights into the mechanisms underlying the relationship of interest. Moreover, we exploit information on the mother's nonstandard work schedules available in the German Microcensus data, including e.g. working on Sundays or at nights. Focusing on this information we investigate whether it matters *when* the mother works, although in a purely descriptive manner (OLS model).

Our results suggest that maternal full-time employment is positively related to child overweight in Germany. OLS estimates indicate that maternal full-time employment is associated with a 5 percentage point higher probability for children aged 9–12 to be overweight. However, when the potential endogeneity of maternal employment is taken into account, estimates become even larger suggesting that children of full-time employed mothers face a 25 percentage point higher overweight probability. Although there are limitations to our IV approach, several robustness analyses point to the validity of our results and the instrument used. The estimates are quantitatively comparable across different specifications and the data used. When we apply the IV approach to explore potential mechanisms we find that maternal employment likely induces a higher consumption of processed food and soda drinks, a lower consumption of vegetables and fruits and an increased sedentary behavior (watching TV and playing video games). Thus, our findings suggest that unhealthy dietary and activity habits might explain the effect of maternal employment on child overweight to some extent. Furthermore, OLS models provide evidence that children of mothers who work nonstandard hours, such as working on Sundays or at nights, face a significant higher overweight probability. Overall, our findings indicate that preadolescent childhood overweight in Germany could be addressed by public policies helping children to live healthy lifestyles and promoting family-friendly employment policies.

The remainder of the chapter is structured as follows: Section 2.2 briefly summarizes the previous literature and discusses some theoretical considerations concerning the underlying mechanisms. Section 2.3 introduces the two data sources, the German Microcensus and the KiGGS data. In Section 2.4 we discuss the empirical strategy used

[7] In analyses not shown, we also considered the amount and quality of sleep, special mechanisms proposed by Ziol-Guest (2014). However, these variables turned out to be not significantly related to maternal employment. As they might thus unlikely serve as mechanisms, we skipped these results due to lack of space.

and the validity of the instrument while Section 2.5 shows the results of these analyses including several pathway and robustness analyses. Section 2.6 concludes.

2.2 Background

One problem that arises when estimating the effect of maternal employment on child outcomes is that maternal employment is likely endogenous for several reasons (Anderson et al., 2003; von Hinke Kessler Scholder, 2008; Greve, 2011; Morrill, 2011; Cawley and Liu, 2012). First, omitted variables might bias the results when certain unobservable characteristics, like ability or effort, motivate some mothers to work.[8] Mothers with such characteristics, such as a high ability or great ambition might be more likely to work but they might also be more able to ensure a healthier lifestyle for the child at the same time, which might in turn be related to a lower overweight probability of the child. For instance, Aughinbaugh and Gittleman (2004) and Ruhm (2004) show that non-employed mothers have on average lower AFQT scores and are less educated compared to employed mothers. Disregarding these characteristics like maternal ability or effort would lead to an underestimation of the OLS estimate. In contrast, women who particularly enjoy spending time with their children might be less likely to work more hours (Anderson et al., 2003; Cawley and Liu, 2012). If those mothers were also more likely to indulge their children, omitting this characteristic would probably result in a downward bias. However, given that the opposite relation is also conceivable, the direction of the related bias is a priori unclear. Reverse causality constitutes a second source of potential bias. For instance, mothers could be more likely to reduce their working hours or stop working if their children suffer from adverse health conditions as these children might need more assistance. Ruhm (2008) finds a significant correlation between child weight and maternal employment measured in the following year. This result provides some evidence that the relationship might be evoked by reverse causality.

In several previous studies, researchers have already tried to tackle these problems and isolate the causal effect of maternal employment on child overweight. Some studies apply child or family fixed effects in order to eliminate time-invariant unobserved child and family heterogeneity (Anderson et al., 2003; Ruhm, 2008; von Hinke Kessler Scholder, 2008; Ziol-Guest et al., 2013). For instance, Anderson et al. (2003) estimate different fixed effects models for children aged 3–11 from the NLSY and find that 10 additional working hours per week increase the child's probability to become overweight by 1.5 percentage points. Other studies use measures of the local labor market or child care conditions as instrumental variables in order to isolate the causal effect of maternal employment on

[8]Of course, the mother's choice to work may be driven by financial distress which depends on the household's financial conditions. We try to address this issue by controlling for the household income and the spouse's employment status.

child overweight (e.g. Anderson et al., 2003; Garcia et al., 2006; Greve, 2011; Datar et al., 2014).

Given the different stages of child development as well as the different age-related child care arrangements available, maternal employment might promote or prevent child overweight at different ages to a different extent. Indeed, previous studies have found that the timing of maternal employment is important, showing that it especially increases the probability of overweight in middle childhood (von Hinke Kessler Scholder, 2008; Miller, 2011; Ziol-Guest, 2014). However, many previous studies focus on a long time span within the childhood and thus estimate an aggregate effect. This might result in misleading conclusions in case that maternal employment promotes child overweight in one certain developmental period while it prevents child overweight in another. Regarding Germany, after-school care has mainly been informal until the mid-2000s, with care provided by relatives playing an important role (see Wrohlich, 2011 for an overview).[9] Given this lack of formal daycare for school-aged children, it remains unclear what happens to the children aged 9–12 when the mothers are working. Being home alone might foster unhealthy dietary habits or sedentary behavior (e.g. media consumption).

Besides the investigation of the mere relationship between maternal employment and child weight, previous literature has increasingly focused on potential mechanisms why maternal employment is related to child overweight. Basically, maternal employment determines the mother's allocation of time spent working and time spent with the child. As mechanisms are most likely related to these time constraints, it is helpful to focus on time-use information, either of the working mothers (and fathers) or of the child itself, to detect potential pathways.[10] Additionally, one could imagine that the time constraints maternal employment involves, lead to less time spent with activities that likely promote child health, such as ensuring a healthy diet. Indeed, some previous studies have found that full-time employed mothers spend less time on food preparation, report fewer family meals and lower levels of fruit or vegetable intake (US: Cawley and Liu, 2012; Bauer et al., 2012; Germany: Möser et al., 2012). For the US, Datar et al. (2014) have shown that unhealthy food, which is more often consumed by children of working mothers, seems to mediate the effect of maternal employment on child overweight. In contrast, Gwozdz et al. (2013) have found little evidence that maternal employment is related to caloric intake and physical activity in selected European regions. Apart from the negative effect of maternal employment on dietary habits, maternal employment could also have a beneficial effect. For instance, a healthier but more expensive nutrition becomes affordable due to the increase in family income. However, previous studies

[9]Daycare slots for children aged 3–6 are statutorily regulated since 1996 leading to the vast majority (90 %) of these children attending daycare (Kindergarten). However, for children below or above this age range, the supply of daycare provisions is comparably low. In Germany, school attendance becomes compulsory at the age of six (see e.g. Wrohlich, 2011; Spiess and Wrohlich, 2008).

[10]Another pathway that is not directly related to the time constraints are food expenditures.

have found that the time constraints seem to dominate the income effect of maternal employment (e.g. Datar et al., 2014). Another conceivable pathway is that maternal employment likely involves an extended period of time without supervision. This might promote an inactive lifestyle of the child with a higher frequency of sedentary behavior, such as watching TV or playing video games. For the US, it has been found that maternal employment is related to increased hours of watching TV (Ziol-Guest et al., 2013; Datar et al., 2014; Fertig et al., 2009). For Europe, there seems to be no relationship (Gwozdz et al., 2013) although this might partly be explained by differences in the supply and quality of child care across the European states. For instance, no relationship between maternal employment and child weight has been found for Denmark which might be due to work and family policies that provide high-quality child care, available healthy food, and family-friendly schedules for both parents (Greve, 2011).[11]

Structures encouraging long work hours and frequent nonstandard schedules (e.g. working on Saturdays/Sundays or in the evening) have been emerging in many developed countries. Although the literature is still scarce, some studies suggest that nonstandard work schedules are also related to child outcomes such as overweight.[12] While Morrissey et al. (2011) find only weak evidence that one summarizing measure of nonstandard work is associated with child overweight in the US, Dunifon et al. (2013) provide some evidence that maternal night work is related to a higher probability of the child to suffer from behavioral problems. Additionally, long work hours by both parents have been associated with child BMI trajectories (Morrissey, 2013).

Briefly, it can be said that evidence on the relationship and its underlying mechanisms how maternal employment is related to child overweight is still inconclusive given the substantial cross-country differences. We are aware of two studies that target this relationship in Germany. The recent study by Gwozdz et al. (2013) focuses on eight different European regions, finding no significant association for Germany. However, as the data cover two small and relatively poor regions in Germany, the results obtained in this study are neither representative nor generalizable to the wider population.[13] Another study by Mahler (2007) uses German SOEP data and finds that maternal full-time employment during childhood increases a child's probability of being obese as a young adult.

[11]Alternatively this may be due to more egalitarian approaches to childrearing and housework which is more common in Scandinavian countries.

[12]See Morrissey et al. (2011) for a detailed discussion on possible links why maternal nonstandard work schedules might be associated with child overweight.

[13]The regions are Delmenhorst and Wilhelmshaven.

2.3 Data and Measures

To estimate the relationship between maternal employment and child overweight, our analyses are based on two sets of representative large-scale surveys from Germany: four pooled cross-sections of the German Microcensus and the cross-sectional KiGGS data.

The German Microcensus Data

Our main data source consists of four cross-sections of the German Microcensus 1999, 2003, 2005, and 2009.[14] These four waves offer the latest available information on health-related questions as these are included every four years. Using the German Microcensus has distinct advantages. First, pooling the four cross-sections leads to a large sample size as the Microcensus is a representative official survey on the living situation of approximately 1 % of German households. Second, we are able to explore several definitions of maternal employment as this data source includes more detailed information on the mother's employment situation. Third, the Microcensus is conducted as a household survey which enables us to determine the age of the youngest child in the household. This is important to perform some robustness checks regarding the validity of the instrument used in this study. The main drawback of the data is that the health-related information in the Microcensus is limited. Moreover, for children aged 15 and younger, information rests upon proxy reports of their parents.[15]

For the purpose of this study, the sample is restricted to children aged 9–12 years living with at least one biological parent. Furthermore, we exclude mothers who are currently enrolled in education. This is, because they might spend less time with their children as attending school leads to additional time constraints. Finally, we focus on children with non-missing data on the variables used in the main specification. The analysis sample size amounts to 45,210 boys and girls aged 9–12 years.

The KiGGS Data

The Microcensus data are restricted due to a lack of proper health-related measures. We thus additionally use the public-use file of the KiGGS data. KiGGS is a nationally representative survey on the health of 17,641 German children aged 0–17 years which

[14] This official data was provided by the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the Länder in Düsseldorf, Germany, analyzed on-site (further information: http://www.forschungsdatenzentrum.de/en/).

[15] In general, it is poorly understood how BMI based on parental reports of body height and weight differs from measured BMI. The limited empirical evidence indicates that parental reports tend to overestimate underweight children and underestimate the overweight children (see Himes, 2009 for an overview). However, as BMI is the outcome, measurement error will only affect the results if the bias differs systematically between non-employed and employed mothers. Moreover, considering the child's overweight status, which dichotomizes the BMI, mitigates the measurement error as the bias within each category becomes negligible.

was conducted during 2003–2006 (see Kurth et al., 2008 for details). Using the KiGGS sample as supplemental data source has several advantages over the German Microcensus. First, apart from collecting data in self-administered parental questionnaires and child questionnaires for children aged 11 years and older, the study included medical interviews and examinations. Hence, body height and weight are measured by trained medical staff. Anthropometric measurements are preferable to reported measures because it is less prone to reporting error e.g. due to social desirability. Second, the data include several child and family characteristics, also at child birth, which allows us to check the robustness of our results. Finally, the data contain detailed information on the child's time use that enables us to explore possible mechanisms. We generally focus on information based on parental assessments for the sake of consistency except for the dependent variable. In rare cases, if only self-assessments are available (e.g. physical activity), we rely on this information for children aged 11 years and older. Restricting the sample analogously to the Microcensus data, the final sample covers 2,447 children (1,283 boys and 1,164 girls).

Measuring Child Overweight

Our main outcome is the child's overweight status (including obese children). We use the age- and sex-specific cut-off point recommended by the international obesity task force (IOT) which is comparable to the commonly used adult cut-off point of a BMI of 25 for being overweight (Cole et al., 2000).[16] In another specification, we use the BMI controlling for completed year-of-age dummies and the child's sex.[17] Within the KiGGS data, we are able to additionally control for the exact age in days. In the samples used for the analyses, 23 % of the children in the Microcensus data and 22 % of the children in the KiGGS data are classified as being overweight.[18]

Measuring Maternal Employment

The main predictor is the mother's full-time employment status collapsing part-time and not employed mothers in the reference category. While full-time employment is defined as usually working more than 32 hours per week in the Microcensus, parents had to self-assess their current employment situation on the basis of given, unspecified categories in the KiGGS data. Especially with respect to part-time employment, the category is rather imprecise capturing part-time as well as hourly employed mothers. Detailed analyses based on maternal working hours reveal the particular importance of maternal full-time employment (see Table 2.A.3 in the Appendix) and part-time employment seems

[16]There is also a German cut-off point for child overweight (Kromeyer-Hauschild et al., 2001). The results are very similar when this definition of overweight is used in the analyses.

[17]Results based on BMI z-scores for children according to US or UK growth rates are very similar.

[18]Table 2.A.4 in the Appendix provides some quantile regression estimates of maternal employment at different points in the BMI distribution.

to be indistinguishable from non-employment (see Table 2.A.2 in the Appendix). For that reason, we confine ourselves to this definition of maternal employment.

Due to the cross-sectional design of both studies, it is not possible to consider the employment history or the cumulative maternal working hours since the child's birth. However, Anderson et al. (2003) find that rather the hours per week than the weeks worked since child birth are directly related to the daily time constraints relating to child overweight. Given the data available, we thus focus on current maternal employment. About 17 % of the mothers in the Microcensus data and 18 % of the mothers in the KiGGS data are full-time employed.

Control Variables

In order to obtain comparable results, we follow previous studies (e.g. von Hinke Kessler Scholder, 2008; Gwozdz et al., 2013) and include three different sets of control variables in our analyses. We distinguish between child characteristics, maternal characteristics and other family characteristics. Child characteristics include the sex, age dummies and a residential variable for living in West/East Germany. In analyses based on the KiGGS data, we are able to include additional controls: the exact age in days and three dummies for different levels of residence urbanization. Moreover, we are also able to control for characteristics of the child's initial health status by including the child's birth weight and an indicator whether the child was breastfed. We also control for the person completing the parental questionnaires to take different patterns of reporting behavior between mothers and fathers into account (cf. Reinhold and Jürges, 2012). Within the Microcensus, the information is exclusively given by the child's mother given that the mother is living in the household. To capture possible time trends we additionally control for the survey year in the analyses based on the Microcensus.

Maternal characteristics include the mother's age group and four dummies capturing the mother's educational level (low, intermediate and high) including a category for missing information. In the analyses based on the KiGGS data, we additionally consider dummy variables indicating whether or not the mother gained weight, smoked or drunk during pregnancy. As the Microcensus lacks such information, we are only able to control for the mother's current smoking status in order to approximate her smoking behavior during pregnancy.

As family characteristics we consider dummies for the number of individuals in the household and two dummy variables whether the father and mother hold the German citizenship control for the child's ethnic background. The mother's employment status is likely influenced by the father's income and time investments which might in turn also be related to the child's weight. The data lack information on the father's income.[19] Although household income is endogenous as it comprises the mother's earnings, we include dummies for the household income as well as the father's employment status to mitigate any potential bias. Moreover, controlling for household income enables us to partly eliminate the effect resulting from higher family income that would otherwise be captured by maternal employment.

We also include the mother's and father's BMI with additional variables that indicate missing information. In some previous studies parental BMI is explicitly excluded from the analyses when estimating the effect of maternal employment on child overweight as it is likely endogenous (e.g. von Hinke Kessler Scholder, 2008). That is because maternal employment might not only affect the child's but also the parent's weight as they likely share the same lifestyle and dietary habits within the family. However, we decided to include parental BMI with an additional category for missing values in our analyses for two reasons. First, parental BMI can also be interpreted as a proxy for genetic factors, which might bias the results as we are unable to control for that. Second, controlling for parental weight status is important with regard to our IV approach. As it might be related to both our instrument and the outcome, controlling for parental BMI is crucial in order to obtain consistent estimates.

Table 2.A.1 in the Appendix summarizes the main variables for both samples and reports the mean differences between overweight and not overweight children.

Mechanisms

In the selection of potential mechanisms that might explain the relationship between maternal employment and child overweight, we mainly follow Datar et al. (2014) and distinguish between different measures for the child's dietary behavior (consumption of vegetables[20]/fruits/soda drinks/sweets per day and consumption of junk food/processed food per week), frequency of physical activity per week and hours of sedentary behavior (watching TV/playing video games per day[21]). The data lack direct information on whether the food is eaten at home or eaten out. Nevertheless, we assume that these measures are comprehensive as dietary patterns were collected by a food diary jointly kept by the parents with their children. The cut-off points regarding the child's diet are generally based on definitions used in previous studies, e.g. by Datar et al. (2014). In some cases, the thresholds deviate from the guidelines recommended by the German

[19] For the Microcensus data, we are able to determine the father's income if the father is living in the same household. The results are very similar when the father's instead of the household income is used or household income is excluded from the analyses (see Table 2.A.7 in the Appendix).

[20] We use a combined measure of eating raw and cooked vegetables.

[21] We use a combined measure of weekdays and weekend days.

Nutrition Society (e.g. five servings of vegetables and fruits per day) as the proportion of children matching these guidelines is very low.[22]

These aforementioned child-related potential mechanisms are available in the KiGGS data only. But the Microcensus data include information on maternal engagement in so-called nonstandard work, i.e. working on Saturdays, working on Sundays or holidays, working in the evening (6 pm–11 pm) or at night (11 pm–6 am), and working at home. Focusing on the different maternal work schedules, we aim to assess whether it matters *when* the mother works.

2.4 Empirical Approach

In order to empirically assess whether maternal employment is related to child overweight and child BMI in Germany, we apply a basic ordinary least squares model (OLS) in a first step. We thus estimate a linear probability model when overweight is the outcome of interest. As previously discussed, relying exclusively on an OLS approach is problematic as the mother's choice to work might be driven by unobserved factors that are also related to the child's probability to be overweight. Previous studies have implemented different econometric approaches in order to isolate the causal effect of maternal employment on child outcomes. Household or child fixed effects (e.g. applied by Anderson et al., 2003; von Hinke Kessler Scholder, 2008; Morrissey et al., 2011) capture unobserved heterogeneity that is constant over time. However, this approach still does not solve the problem if the relationship is driven by reverse causality or if changes in maternal employment are not exogenous to the outcome. For instance, mothers might reduce their working hours to supervise their overweight children. Another commonly used approach is to find an instrumental variable that is correlated with maternal employment but uncorrelated with the error term. The difficulty is to find strong instruments that fulfill this validity assumption. In previous studies maternal employment is most often instrumented by local labor market conditions (Anderson et al., 2003; Greve, 2011; Bishop, 2011; Datar et al., 2014). However, although such instruments seem exogenous at first sight, Cawley and Liu (2012) raise concern about the validity of such instruments. For instance, there is some evidence that macroeconomic factors, such as the local unemployment rate, affect health (e.g. Ruhm, 2006) and thus also the outcome of interest. In this study, we also try to address the possible endogeneity of maternal employment by a two stage least

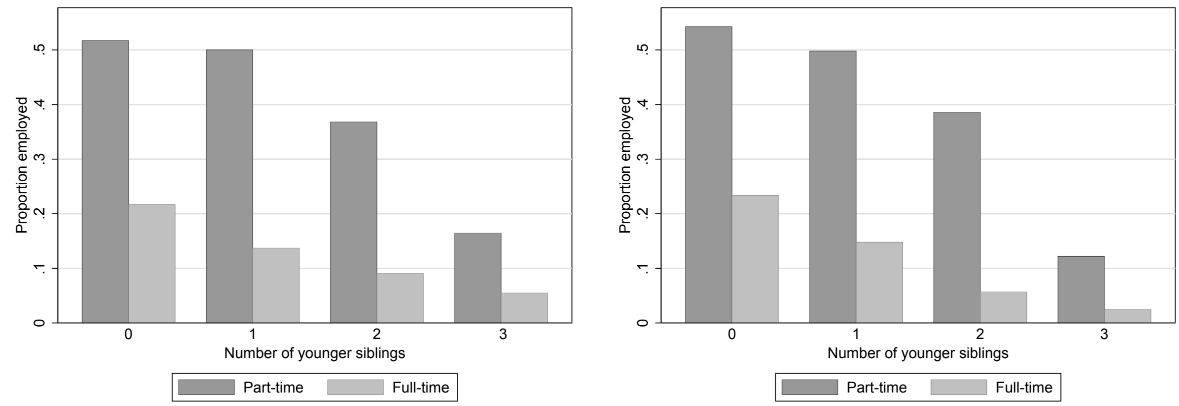
[22] For some mechanisms (e.g. watching TV or playing video games) the cut-off points deviate from the recommended guidelines as these variables are collected in fixed categories that do not match with the definitions used by Datar et al. (2014). Robustness analyses using different thresholds point to the same direction and are available upon request.

squares approach (2SLS) and instrument maternal full-time employment by the number of younger children in the household.[23]

2.4.1 The Instrument: Number of Younger Siblings

Taking the number of younger siblings in the household as instrument rests upon the idea that the mother's incentive to work (full-time) is considerably lower if there is at least one younger child in the household. Regarding the mother's allocation of time, the opportunity costs to stay at home and care for the child of interest (older sibling) is substantially lower when there is another, younger child she has to care for already. Figure 2.1 shows the relationship between the number of younger siblings in the household and maternal employment status within the two data sets. In general, it points to the expected negative relationship: The share of part-time and full-time employed mothers decreases with the number of younger siblings. While the highest share of full-time employed mothers is among those without younger children (Microcensus: 21 %, KiGGS: 23 %), the lowest can be found among mothers with at least three younger children in the household (less than 5 %). The descriptive relationship thus suggests that the number of younger siblings is highly relevant for maternal (full-time) employment which will be tested in the first stage regression (*instrument relevance*).

Figure 2.1: Maternal employment by the number of younger siblings



(a) *Source*: German Microcensus 1999, 2003, 2005, 2009.

(b) *Source*: KiGGS 2003–2006.

Another requirement in order to obtain valid IV estimates is that the instrument must be uncorrelated with the error term (*instrument exogeneity*). That is, conditional on the variables included in the model, the only path through which the number of younger siblings in the household is related to the child's overweight status/BMI is maternal employment. Theoretically, one might expect that each additional (younger) child involves

[23]See Angrist (2001) and Morrill (2011), for a discussion on 2SLS models with a binary dependent variable.

further time constraints which impair the maternal care for the older child, e.g. to cook healthy meals. In this case, the instrument might be invalid, as it likely correlates with the child's weight. However, cooking succumbs substantial economies of scale which implicate that cooking for one child is more or less equal to the effort required to cook for two or three children. More generally, one might suspect that the number of younger siblings as a family characteristic is related to factors that determine the child's overweight status, e.g. the father's time investments, the financial condition of the family, genes or shared dietary habits. To address this, we include the father's employment status, family income and the parental BMI. That way, we aim to capture these channels through which the instrument could be related to the child's weight status and thus mitigate any potential bias.

Other concerns which are even more difficult to address arise from the fact that the number of younger siblings in the household might itself be endogenous and correlated with unobservable characteristics, like maternal ambition or effort (c.f. Morrill, 2011). For instance, one could argue that more ambitious mothers are more likely to work full-time but are in turn also more likely to have more (younger) children. We argue that this bias is negligible as the number of children generally decreases with female educational level (Destatis, 2012) which is in turn likely positively correlated with ambition. If ambitious mothers were nevertheless more likely to have more younger children, the IV estimates might be downward-biased. The downward bias results from the positive correlation between maternal ambition and the probability to work on the one hand, but the negative correlation between maternal ambition and the child's probability to be overweight on the other hand. As maternal full-time employment appears to be strongly positively correlated with the child's BMI and overweight probability in this study, this bias would likely result in an underestimation of the true effect.

Another concern is related to the fertility choice that determines the number of younger siblings. Families with multiple children might be different from families with single children. For instance, an additional younger child might affect the weight status of the older child. That is because an additional child likely changes the parental behavior and limits the parents' time, including the time spent in physical activity and food preparation. A priori it is unclear in which direction this would bias our results. On the one hand, parents of multiple children might be more permissive with regard to unhealthy food or dietary habits which may bias the results upwards. On the other hand, parents might be more engaged to ensure healthy food and living if they have multiple children. In this case, the child would be less likely to become overweight resulting in a downward bias. Recent studies have shown that singleton children are more likely to be overweight than their peers with siblings (e.g. Hunsberger et al., 2012), lending support for the downward bias. In addition, we argue that controlling for parental BMI captures dietary and activity habits within the family at least to some extent.

Selective fertility might still impair the validity of our instrument, e.g. through an unobserved family-orientation. For instance, family-orientated mothers might be more likely to have multiple children and care for them but are less likely to work. In order to explore whether our results could be driven by selective fertility, we perform a robustness analysis restricting the sample to families with two children (of different ages). Although we are still not able to investigate how fertility choice affects our results in a comprehensive way, we are able to assess the effect for families with the same number of children and thus with an equal fertility choice at the time of the survey. With respect to our instrument, we thus compare families where the child of interest has a younger sibling (instrument=1) with children having an older sibling (instrument=0).

Relatedly, the number of younger siblings correlates with the birth order of the children. We are not aware of any study assessing the link between being first-born (or birth order in general) and child overweight in a causal sense. But few epidemiological studies found that first-born children are more likely to become overweight (e.g. Li et al., 2007) which would impair the validity of our IV estimates. Given the nature of the phenomenon, it is arguably difficult to identify a causal link in this context, as the appropriate reference group for first-born children consists of later-born children without older siblings. Moreover, this correlation might occur exactly because of maternal employment, a social mechanism probably neglected by the epidemiological literature. Nevertheless, in a robustness check in Section 2.5.2, we exclude first-born children and children without older siblings in the household, in order to eliminate the effect arising from being the first-born child. That way we are at least able to check whether the IV estimates are biased due to a systematic correlation between being first-born and child overweight.

We believe that while controlling for a range of family and child characteristics to mitigate any potential bias, the number of younger siblings is a good instrument as it is strongly related to maternal full-time employment. Hence, when discussing the IV results we interpret the 2SLS estimates as effects although this interpretation rests upon the assumption that the instrument is valid. We treat the instrument linearly in the main specifications as this definition has the strongest predictive power. In a robustness check, we also include the number of younger siblings as dummy variables.[24]

2.4.2 Mechanisms

We explore several possible mechanisms why maternal employment is related to the child's overweight status. Given the problems that arise in identifying mediators in a causal sense (see Bullock et al., 2010) we do not perform a full mediation analysis. Instead, we examine whether maternal employment is causally related to the possible mediation variable and

[24]In a robustness analyses not shown, we also excluded families without younger siblings in the household. As the results appear to be robust, we focus on instrumenting maternal full-time employment by the continuous measure of the number of younger siblings in the household in all specifications.

thus only estimate the first path in the Baron and Kenny (1986) mediation framework. Any variable that is significantly related to maternal employment indicates a possible channel through which maternal employment is related to child weight. However, as we are not able to test the other path, the results cannot be interpreted as causal evidence for a mechanism in the relationship of interest. In general, we thus follow the strategy used by Cawley and Liu (2012) to identify mechanisms that could potentially be responsible for the effect of maternal employment on child overweight. In contrast to Cawley and Liu, we focus on the child's instead of the mother's time-use data. Moreover, we aim to estimate the effect of maternal full-time employment on the specific mechanisms using our IV approach.

Based on the Micorcensus data, we assess whether it matters *when* the mother works, exploring nonstandard work schedules. In contrast to the investigation of the other mechanisms, we rely on the descriptive results obtained from simple OLS models as the instrument used is not predictive for these definitions of maternal nonstandard employment.

2.5 Results

The main analyses are built on both data sources. We start reporting the results based on the German Microcensus (Panel A) as it comprises a much larger sample size leading to a higher precision of the estimates. In a second panel we report the findings based on the KiGGS data (Panel B) in order to assess whether these estimates point to similar conclusions when a broader range of child and family characteristics is controlled for that are unavailable in the Microcensus. All models include a full set of control variables in order to minimize any potential bias as discussed in Section 2.3.[25] The first row of each panel presents the results for the continuous BMI variable while the second row presents the results for the overweight status indicator variable. The mean BMI and overweight prevalence are included to better assess the dimension of the estimated coefficients.

2.5.1 Main Results

Table 2.1 reports the main results of the relationship between maternal employment and child BMI/overweight status for both data sources. We first discuss the OLS results presented in Column (1). In the Microcensus (Panel A) maternal full-time employment is significantly related to child BMI: The BMI of children with full-time working mothers is on average 0.38 kg/m^2 higher compared to children of not employed or part-time employed mothers. Accordingly, maternal full-time employment is associated with a 3.5

[25]Table 2.A.2 in the Appendix presents the OLS estimates when the control variables are included successively, distinguishing between maternal part-time and full-time employment. Table 2.A.3 explores several definitions of maternal employment.

percentage point higher probability of the child being overweight. The OLS results based on the KiGGS (Panel B) reveal a similar pattern: Maternal full-time employment is significantly related to a higher child overweight (4.9 percentage points) and a higher BMI (0.53 kg/m^2). These positive associations are consistent with findings from studies using US or UK data (e.g. von Hinke Kessler Scholder, 2008; Datar et al., 2014) and with the study by Mahler (2007) finding a positive relationship for Germany.

Table 2.1: The effect of maternal full-time employment on child weight

	Mean	(1) OLS	(2) FS	(3) RF	(4) 2SLS	N
A: Microcensus						
BMI	18.3235	0.3765***	-0.0730***	-0.2056***	2.8150***	45,210
	(0.0162)	(0.0454)	(0.0024)	(0.0228)	(0.3235)	
Overweight	0.2324	0.0353***	-0.0730***	-0.0206***	0.2815***	45,210
	(0.0020)	(0.0056)	(0.0024)	(0.0029)	(0.0403)	
B: KiGGS						
BMI	18.4871	0.5285***	-0.0935***	-0.1788*	1.9116*	2,447
	(0.0691)	(0.1683)	(0.0103)	(0.1028)	(1.1012)	
Overweight	0.2150	0.0486**	-0.0935***	-0.0234**	0.2504*	2,447
	(0.0083)	(0.0227)	(0.0103)	(0.0123)	(0.1322)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; FS refers to the first stage, RF to the reduced form and 2SLS to the second stage estimations; the first stage F-statistics are 82.87 for the KiGGS data and 888.04 for the German Microcensus; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

Although the OLS models suggest that maternal full-time employment is positively related to a child's (over)weight, the relationship might still occur due to unobserved characteristics of mothers and other non-causal explanations, which we are unable to control for. The main analysis thus aims to take the endogeneity of maternal full-time employment into account, by applying the IV approach discussed in Section 2.4. The first stage (FS), which is shown in Column (2), regresses maternal full-time employment on the instrument. The results indicate that the number of younger siblings in the household is strongly related to maternal full-time employment across the two data sources. The estimates are significant at the 1 % level. Our approach therefore does not suffer from a weak instrument problem. In accordance with the descriptive evidence presented in Figure 2.1, the estimates indicate a negative relationship: When the number of younger siblings increases, the probability that the mother works full-time decreases. The causal interpretation of the IV estimates in the second stage rests upon the non-testable assumption that the instrument affects child (over)weight exclusively through maternal employment, the endogenous variable. We report the estimates of the reduced form (RF), i.e. the regression of the outcome on the instrument, in Column (3). The RF estimates are consistently negative and significantly different from zero, at least at the 10 % level. The coefficient for the Microcensus (KiGGS) suggests that the BMI significantly decreases by 0.21 kg/m^2 (0.18 kg/m^2) with each additional younger sibling in the household. Accordingly, the probability of the child of interest to be overweight decreases by 2.1 (2.3) percentage points. The negative coefficients in both the first stage and reduced form estimation are in line with our theoretical predictions (Section 2.4).

Column (4) shows the second stage estimates, which reflect the effect of maternal full-time employment on child overweight.[26] As expected from the negative coefficients in the reduced form and first stage models, the estimates of the second stage are positive and significant regardless of the data used. In the Microcensus data (Panel A), maternal full-time employment significantly increases child BMI by 2.8 kg/m^2 and the probability to be overweight by 28.2 percentage points, respectively. In relation to the mean value, this corresponds to an increase of about 120 %. To illustrate, taking a 10-year old boy with average BMI and height as an example, the estimated effect would correspond to an increase in body weight of about 4 kg because of his mother's full-time employment. Similarly, the IV estimate obtained from the KiGGS (Panel B) points to a 25 percentage point higher overweight probability due to maternal employment. Although the 2SLS results based on the KiGGS data are less precise and significant at the 10 % level only, we interpret this finding as evidence for the robustness of our results, as the conclusions drawn from both data sources are very similar.

Overall, the findings are again in line with previous studies using US and UK data finding slightly larger estimates for maternal employment when taking the possible endogeneity of maternal employment into account (Anderson et al., 2003; von Hinke Kessler Scholder, 2008; Bishop, 2011). This suggests that mothers may select themselves into full-time employment due to some unobserved characteristics that promote child health, like effort or ability, resulting in lower estimates obtained from OLS models. Another reason for this downward bias could be that maternal full-time employment is not measured accurately. Commuting hours or other time constraints that maternal full-time employment involves might not be captured by the contractually agreed employment definition, as reported in the data used. However, the estimates obtained from our IV approach become large in magnitude relative to the OLS coefficients. It has to be considered that the instrumental variable approach estimates a local average treatment effect which allows to interpret the identified effects only for the narrow subpopulation of compliers (Angrist and Pischke, 2009). In our case, the compliant population consists

[26]See also Table 2.A.5 in the Appendix for different estimation strategies of the IV model. The estimates in Column (6) are derived using a two-stage-residual-inclusion approach in which the second stage includes the treatment and the fitted residual from a probit-estimated first stage in order to take the binary nature of the treatment into account (Wooldridge 2010, pp. 126–129; Terza et al. 2008).

of mothers whose decision to work was influenced by the number of younger siblings in the household. These mothers might have certain characteristics or, for instance, make a huge effort that accounts for the large estimates. Obtaining large IV estimates is not unusual. Morrill (2011) also finds that estimates increase substantially when applying a similar IV approach using the younger sibling's kindergarten eligibility as an instrument for maternal employment. Although the estimates are smaller and not directly comparable, as Morrill focuses on different child health outcomes other than overweight (e.g. the probability of an overnight hospitalization) this still supports the credibility of our estimates. The increase in the effect size might also result from an invalid instrument, i.e. the instrument correlates with factors that are also related to the child's weight status. For that reason, we provide some robustness analyses addressing the concerns discussed in Section 2.4.1.

2.5.2 Robustness Analyses

We next test the robustness of our estimates and the instrument used. Table 2.2 presents the results of the OLS model, the first stage (FS), the reduced form (RF) and the 2SLS estimates for each specification.

We first exploit the Microcensus data in order to check the robustness of our instrumental variable approach. Within the Microcensus, we are able to determine the younger siblings' age which might be crucial for the mother's employment decision. A growing stream of literature finds that the availability of low-cost or free public child care increases maternal employment among mothers without younger children (e.g. Cascio, 2009; Fitzpatrick, 2010). Building on this literature, we define our instrument by the number of younger siblings below the age of six as school attendance becomes compulsory at that age in Germany. As we base these analyses on the same sample of children, the coefficients obtained from the OLS models (Column 1) are equal to those of the main specification presented in Table 2.1. In line with the results previously obtained, we again find that the instrument is highly predictive for maternal employment. The results obtained from the second stage are quantitatively and qualitatively very similar to the main specification.

In a related instrument definition presented in Panel (2), we include dummies by age group distinguishing between the following categories: no younger sibling in the household (reference group), at least one younger sibling aged 3 years or younger in the household (no institution, mother most likely in maternity leave), at least one younger sibling aged 3–6 years in the household (Kindergarten) and at least one younger sibling aged 6 years or younger in the household (school age). In the first stage regression (see Table 2.A.6) maternal full-time employment is significantly and negatively related to all dummies included in the analyses. As expected from the theoretical discussion,

Table 2.2: Robustness analyses

Model specification	Dependent	Mean	(1) OLS	(2) FS	(3) RF	(4) 2SLS	N
A: Microcensus							
(1) Alternative Instrument: # younger siblings <6 years of age							
	BMI	18.3235	0.3765***	-0.0635***	-0.1625***	2.5579***	45,210
		(0.0162)	(0.0454)	(0.0036)	(0.0347)	(0.5578)	
	Overweight	0.2324	0.0353***	-0.0635***	-0.0182***	0.2862***	45,210
		(0.0020)	(0.0056)	(0.0036)	(0.0042)	(0.0677)	
(2) Alternative Instrument: # younger siblings, dummies by age group (see 2.A.6 for FS and RF)							
	BMI	18.3235	0.3765***			2.9949***	45,210
		(0.0162)	(0.0454)			(0.3304)	
	Overweight	0.2324	0.0353***			0.2811***	45,210
		(0.0020)	(0.0056)			(0.0041)	
(3) FT vs. PT: Not employed mothers excluded							
	BMI	18.23573	0.3914***	-0.0628***	-0.2311***	3.6785***	29,748
		(0.0024)	(0.0485)	(0.0040)	(0.0313)	(0.5425)	
	Overweight	0.2193	0.0385***	-0.0628***	-0.0204***	0.3248***	29,748
		(0.024)	(0.0061)	(0.0040)	(0.0040)	(0.0659)	
(4) First-born and children without older siblings excluded							
	BMI	18.4485	0.3394***	-0.0616***	-0.1221***	1.9819***	22,367
		(0.0225)	(0.0638)	(0.0038)	(0.0356)	(0.5852)	
	Overweight	0.2475	0.0389***	-0.0616***	-0.0152***	0.2462***	22,367
		(0.0029)	(0.0083)	(0.0038)	(0.0046)	(0.0761)	
(5) Fertility choice: Restricted to families with 2 children (of different ages)							
	BMI	18.2187	0.3725***	-0.0385**	-0.0959	2.4908	22,999
		(0.02237)	(0.0675)	(0.0152)	(0.1571)	(4.1542)	
	Overweight	0.2231	0.0355***	-0.0385**	-0.0116	0.3018	22,999
		(0.0027)	(0.0081)	(0.0152)	(0.0188)	(0.0677)	
B: KiGGS							
(6) Controlling for child's chronic conditions							
	BMI	18.4501	0.5283***	-0.0894***	-0.1844***	2.0613*	2,336
		(0.0702)	(0.1709)	(0.0103)	(0.1045)	(1.1747)	
	Overweight	0.2098***	0.0490**	-0.0894***	-0.0204*	0.2279*	2,336
		(0.0084)	(0.0231)	(0.0103)	(0.0123)	(0.1378)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; FS refers to the first stage, RF to the reduced form and 2SLS to the second stage estimations; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

the strongest relationship can be found for children with at least one very young sibling (< 3 years) compared to children without younger siblings. The second stage results are again similar to the results of the instrument specification used in the main analyses. We conclude that the results are not sensitive to the definition of the instrument.

In a next step we perform several subgroup analyses in order to assess the validity of the instrument introduced addressing the concerns discussed in Section 2.4.1. Based on restricted samples according to different selection criteria, we repeat the analyses of the main IV specification presented in Table 2.1. In order to obtain estimates of higher precision, we again focus on the German Microcensus as it comprises a much larger number of children. In Panel (3), we exclude children with non-working mothers from our analyses. Therefore, we focus on the variation in hours worked full-time versus part-time. This enables us to estimate the effect of maternal employment on child overweight at the intensive margin and to mitigate the potential remaining bias resulting from the selection into employment. The OLS estimates for BMI (0.39) and overweight (0.039) are still significant and very similar compared to the non-restricted analyses. The results of the first stage are slightly smaller but still significant, indicating that each additional younger child in the household decreases the employed mother's probability to work full-time by 6.3 percentage points. Compared to the main analyses presented in Section 2.5.1, the estimates of the second stage are somewhat larger although qualitatively and quantitatively still similar. Therefore, the IV results of the main specification are unlikely affected by a (remaining) selection into employment.

In order to assess whether the IV estimates of maternal employment on child overweight are driven by the fact that first-born children are at an increased risk to become overweight, we exclude first-born children and children without older siblings in the household in Panel (4). The estimates become slightly smaller but are again quantitatively and qualitatively comparable to the estimates in the main analyses. It is thus unlikely that our results are driven by the correlation between being the first-born child and overweight.

To further explore whether the fertility choice or an unobserved family orientation might be a channel that biases our IV estimates, we restrict the sample to families with two children (of different ages). Thus, we condition on the same fertility choice at the time of the survey in Panel (5). Although the results of the first stage regression still point to a negative relationship, the number of younger siblings seems to be less predictive for maternal employment in two-child families as the estimates become smaller and less precise. This might be due to the smaller variation of the instrument given the restriction on two-child families. The standard errors of the RF and 2SLS regressions become large and the coefficients are not significantly different from zero. Quantitatively the coefficients are still comparable to the estimates obtained from the main estimations in Table 2.1. We conclude that the main results are unlikely driven by differences in fertility although we cannot fully rule out this channel.

In a final robustness analyses presented in Panel (6), we aim to assess whether the observed relationship might be driven by child health problems occurring after birth. As response to a child's chronic condition the mother might stop working (or reduce working hours). In turn, child health might also be correlated with the child's overweight status. Based on the KiGGS data we are able to include variables indicating whether or not the child currently suffers from a chronic condition (although chronic conditions might be endogenous to child overweight) to approximate conditions occurring prior to the mother's employment. Overall, the results are again very similar compared to the main results based on the KiGGS data presented in Table 2.1.

2.5.3 Mechanisms

We further explore which mechanisms might be responsible for the positive effect of maternal full-time employment on child (over)weight. For the KiGGS data, Table 2.3 presents the OLS and IV estimates of the effect of maternal full-time employment on different outcomes. These outcomes might theoretically serve as mechanisms why maternal employment affects child over(weight). Each row presents a regression with another dependent variable. Concerning dietary habits, the results are twofold. Whereas we do not find any significant relation between maternal full-time employment and eating sweets or chocolate at least once a day, or eating junk food at least once a week, we do find some evidence regarding other dietary habits. Although not indicated by the OLS models, maternal full-time employment significantly reduces the child's probability to eat vegetables and fruits at least once a day by 58 and 41 percentage points. Moreover, children with full-time working mothers have a 42 percentage point higher likelihood to drink sugary soda at least once a day. They are also more likely (32 percentage points) to consume processed food once a week compared to children of not employed or part-time employed mothers. Our findings indicate that children of full-time employed mothers consume healthy and freshly prepared food less often. This is in line with previous literature exploring the mechanisms, finding that maternal employment is related to a child's unhealthy dietary habits in the US (Bauer et al., 2012; Datar et al., 2014).

With respect to the child's activities we find no significant difference in the probability to exercise at least three times a week between children of full-time employed mothers on the one hand and children of part-time or not employed mothers on the other hand. While the IV estimate suggests that children of full-time employed mothers are less likely to exercise at least three times a week, this turns out to be insignificant at conventional levels. However, we find some evidence that children of full-time employed mothers are significantly more likely to engage in sedentary behavior. The IV estimate suggests a 47 percentage point higher probability of watching TV at least one hour a day for children of full-time employed mothers. With regard to the probability of playing video games at

Table 2.3: The effect of maternal full-time employment on possible mechanisms

Dependent Variable	Mean	OLS	2SLS	N
B: KiGGS				
Dietary habits				
Vegetables at least once per day	0.2904	0.0024	-0.5754***	2,359
	(0.0093)	(0.0261)	(0.1850)	
Fruits at least once per day	0.5286	0.0176	-0.4135**	2,359
	(0.0102)	(0.0273)	(0.1824)	
Soda drinks at least once per day	0.3157	0.0629**	0.4229***	2,376
	(0.0095)	(0.0268)	(0.1572)	
Sweets at least once per day	0.1987	-0.0311	-0.0502	2,371
	(0.0082)	(0.0215)	(0.1460)	
Chocolate at least once per day	0.1490	0.0090	0.1546	2,369
	(0.0073)	(0.0204)	(0.1285)	
Junkfood at least once per week	0.1173	0.0455**	-0.0041	2,369
	(0.0066)	(0.0197)	(0.1156)	
Processed food at least once per week	0.1529	0.0051	0.3194**	2,354
	(0.0074)	(0.0210)	(0.1300)	
Activity behavior				
Exercise at least 3 times per week	0.4217	0.0006	-0.1313	2,426
	(0.0100)	(0.0248)	(0.1549)	
Watching TV at least 1 h/day	0.9253	0.0333**	0.4747***	2,395
	(0.0054)	(0.0138)	(0.1249)	
Video games at least 1 h/day	0.3893	0.0830***	0.4873***	2,371
	(0.0100)	(0.0273)	(0.1628)	

Note: Each coefficient belongs to a separate regression with the corresponding mechanism variable as dependent; All models include a full range of child, mother and family controls as well as parental BMI; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* KiGGS 2003–2006.

least 1 hour a day, we find a significant 49 percentage point, and thus a 125 %, increase due to maternal full-time employment. These findings are again in line with previous studies finding that maternal employment is related to a higher TV consumption of children at a similar age (Fertig et al., 2009; Ziol-Guest et al., 2013; Datar et al., 2014).

Based on the Microcensus data, we finally assess whether it matters *when* the mother works. We explore nonstandard work schedules relying on simple OLS models. Table 2.4 reports the results on the relationship between different maternal nonstandard work schedules and child overweight/BMI. Children of mothers who work on Saturdays at least sometimes face a significantly higher overweight probability of 2.4 percentage points. Working on Sundays, at nights and in the evening point to a very similar significant relationship, although these correlations appear to be somewhat weaker. Interestingly, we do not find a significant relationship between maternal teleworking and child overweight. This is in line with Rapoport and Le Bourdais (2008) who have previously shown that working at home is in general related to increased time parents spent with different types of domestic work such as cooking and shopping.

Table 2.4: OLS estimates: Maternal nonstandard work and child weight

		Dependent		
Maternal nonstandard work	Mean	BMI	Overweight	N
A: Microcensus				
Work on Saturdays	0.4079	0.1823***	0.0240***	29,971
	(0.0028)	(0.0375)	(0.0048)	
Work on Sundays/Holidays	0.2357	0.1477***	0.0157***	29,970
	(0.0025)	(0.0435)	(0.0056)	
Work at nights (11pm–6am)	0.0913	0.1164*	0.0143*	29,954
	(0.0017)	(0.0644)	(0.0083)	
Work in the evening (6pm–11am)	0.3196	0.0843**	0.0115**	29,962
	(0.0027)	(0.0389)	(0.0050)	
Work at home	0.1481	-0.0483	-0.0033	29,985
	(0.0021)	(0.0504)	(0.0065)	

Note: Nonstandard work = 1 if sometimes, regularly, and often; Nonstandard work = 0 if never; All models include a full range of child, mother and family controls as well as parental BMI; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009.

2.6 Conclusions

The aim of this study was to explore maternal (full-time) employment as a determinant of preadolescent overweight in Germany. In accordance with the studies investigating this relationship for the US (e.g. Ruhm, 2008; Ziol-Guest et al., 2013), we find that children of full-time employed mothers are at an increased risk to be overweight in Germany. OLS models suggest that maternal full-time employment is related to a 5 percentage point higher probability of the child to be overweight. Addressing the endogeneity of maternal employment by an IV approach with the number of younger siblings as an instrument, the estimates indicate a 25 percentage points higher overweight probability due to maternal full-time employment. This suggests that the simple correlation between maternal employment and child overweight likely understates the effect. Mothers may thus select themselves into full-time employment due to unobserved characteristics (e.g. maternal effort, ability) that promote child health. Additional analyses in order to explore the validity of the instrument introduced in this study confirm the robustness of this result.

One strength of our study is that the analyses are built on two different representative data sources from Germany, each having distinct advantages. The results are very similar across the two data sets which supports the credibility of the findings. We focus on children aged 9–12 years, a crucial age range which has yet rarely been studied in the context of maternal employment. Further, we apply our IV approach to explore the role of several dietary and activity habits as mechanisms. The results indicate that maternal full-time employment promotes unhealthy dietary habits, namely a lower consumption of fruits and vegetables, and a higher consumption of soda drinks and processed food. Children of full-time employed mothers also spend more time watching TV and playing video games. This dietary and sedentary behavior might be relevant mechanisms through which maternal employment promotes child overweight. Moreover, descriptive OLS results suggest that mothers' nonstandard work schedules are also related to child overweight.

Our study also has limitations. First, due to the cross-sectional data and the contemporaneous measurement of maternal employment and child BMI, the chronology of treatment and outcome remains unclear. Moreover, current maternal employment unlikely influences the child's weight status directly. However, we find that maternal employment affects the child's dietary and activity behavior, which are more likely to be immediately influenced by maternal employment. Current BMI might thus serve as a kind of proxy for an accumulation of various health behaviors in the past. Second, whereas our results are generally robust across all specifications explored, we cannot fully ensure that unobserved characteristics related to fertility choice (e.g. the mother's family orientation) are properly taken into account by the instrument used in this study. Although the results in a related robustness check are of comparable magnitude, estimates become imprecise. This reflects the trade-off between accuracy and statistical power, especially when applying IV estimations.

Our study nevertheless provides evidence on an undesirable side effect of maternal employment in Germany: preadolescent (over)weight. But the implications of our findings are not that mothers should drop out of the labor force nor that maternal employment is per se harmful for child health. Female (and maternal) employment shares are expected to continue rising in Germany like in other countries as well. Against that background, our results reveal important implications how public policy could be successful in preventing child overweight. Based on the evidence on potential mechanisms found in this study, public policy should invest in ways to facilitate access to or provide healthy food and activities for children and families. In Germany, daycare arrangements for secondary-school-aged children are still limited, although expanding during the last years. Policymakers should encourage schools to incorporate physical activity and healthy meals within daycare to offset the increase in sedentary behavior and unhealthy diet at home. The ongoing public debate on enabling healthier lifestyles for children within daycare and campaigns for healthy school meals might be a step in the right direction. For instance, the campaign "*IN FORM – German national initiative to promote healthy diets and physical activity*"[27] initiated in 2008 by the federal government implements measures within the school environment, such as healthy school meals following the nutritional guidelines recommended by the German Nutrition Society. However, these campaigns rest on voluntary participation and parents largely bear the costs of the school meals themselves. Future research is thus necessary to evaluate how successful these campaigns effectively are in reducing child overweight. Our findings with respect to nonstandard work schedules further reveal that the job conditions of working parents might also be a relevant factor in promoting child health. In addition to the implications regarding school-based daycare, child overweight might be prevented by improving the work-life balance of dual-earner families and working single mothers, e.g. by increasing teleworking and flexible working hours. However, as we are only able to provide suggestive evidence, future research should evaluate to what extent family-friendly working conditions are promoting healthier family lifestyles in Germany.

[27]For further information, see `http://www.in-form.de`.

Appendix

Table 2.A.1: Summary statistics of the main variables (mean values)

		A: Microcensus			B: KiGGS		
		Not Overweight	Overweight	Difference	Not Overweight	Overweight	Difference
Predictor							
M' employment	None	0.331	0.379	-0.048***	0.321	0.328	-0.007
	part-time	0.504	0.434	0.071***	0.511	0.462	0.049**
	full-time	0.165	0.187	-0.022***	0.167	0.209	-0.042**
M' working hours/week		22.74	24.09	-1.355***			
Child characteristics							
	Girl	0.507	0.431	0.076***	0.471	0.492	-0.021
	Age	10.54	10.44	0.101***	10.92	11.01	-0.090
	Breastfed				0.806	0.728	0.078***
	Birth weight				3364	3496	-132.6***
	West Germany	0.838	0.814	0.024***	0.697	0.696	-0.001
	Rural				0.239	0.262	-0.023
	Small town				0.286	0.274	0.012
	Town				0.284	0.285	0.001
	City				0.192	0.179	0.013
Interview	Mother				0.913	0.892	0.021
	Father				0.049	0.067	-0.017
	Parents				0.037	0.040	-0.003
Maternal & family characteristics							
	Mother's age	39.29	38.77	0.522***	38.92	38.69	0.226
During pregnancy	Weight gain				13.11	13.50	-0.392
	Smoke[a]	0.254	0.336	-0.081***	0.126	0.217	-0.091***
	Drink				0.146	0.106	0.039**
Father	Employed	0.918	0.878	0.040***	0.924	0.861	0.063***
HHincome	Low	0.270	0.363	-0.093***	0.280	0.380	-0.101***
	Medium	0.417	0.389	0.028***	0.417	0.399	0.018
	High	0.313	0.248	0.065***	0.303	0.221	0.082***
M' education	Low	0.317	0.439	-0.122***	0.205	0.251	-0.052**
	Intermediate	0.425	0.396	0.029***	0.462	0.504	-0.042**
	High	0.253	0.158	-0.095***	0.320	0.209	0.110***
	Missing	0.005	0.007	-0.002**	0.014	0.030	-0.017***
	Household size	2.688	2.785	-0.097***	2.780	2.751	0.029
	Mother German	0.886	0.833	0.053***	0.941	0.890	0.051***
	Father German	0.891	0.836	0.055***	0.927	0.880	0.046***
Parental BMI							
Father	BMI	26.06	27.43	-1.364***	25.95	27.90	-1.948***
	Missing	0.016	0.012	0.004***	0.084	0.116	-0.032**
Mother	BMI	23.77	25.70	-1.933***	24.99	26.97	-2.983***
	Missing	0.035	0.040	-0.004**	0.007	0.010	-0.003
	%	76.76	23.24		78.50	21.50	
	N	34,703	10,507		1,921	526	

Note: [a]Approximated by the mother's current smoking status in Microcensus; M' refers to maternal; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

Table 2.A.2: OLS estimates: Maternal employment and child weight

Dependent		(1)	(2)	(3)	(4)	(5)
A: Microcensus						
BMI	Part-time	-0.3962***	-0.3849***	-0.2084***	-0.0876***	-0.0276
		(0.0364)	(0.0359)	(0.0361)	(0.0366)	(0.0356)
	Full-time	0.1429***	0.0316	0.2224***	0.3503***	0.3598***
		(0.0500)	(0.0517)	(0.0515)	(0.0523)	(0.0509)
Overweight	Part-time	-0.0511***	-0.0448***	-0.0257***	-0.0139***	-0.0066
		(0.0044)	(0.0045)	(0.0045)	(0.0046)	(0.0045)
	Full-time	-0.0017	-0.0048	0.0160**	0.0302***	0.0312***
		(0.0061)	(0.0063)	(0.0063)	(0.0065)	(0.0063)
	Mean BMI	18.3235	18.3235	18.3235	18.3235	18.3235
		(0.0162)	(0.0162)	(0.0162)	(0.0162)	(0.0162)
	Mean overweight	0.2324	0.2324	0.2324	0.2324	0.2324
		(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0020)
	N	45,210	45,210	45,210	45,210	45,210
B: KiGGS						
BMI	Part-time	-0.2370	-0.3024**	-0.2594*	-0.0810	0.0100
		(0.1553)	(0.1483)	(0.1493)	(0.1499)	(0.1426)
	Full-time	0.3892**	0.3530*	0.3897*	0.5158**	0.5345***
		(0.2087)	(0.2046)	(0.2046)	(0.2005)	(0.1893)
Overweight	Part-time	-0.0203	-0.0185	-0.0117	0.0017	0.0097
		(0.0186)	(0.0186)	(0.0186)	(0.0190)	(0.0186)
	Full-time	0.0365	0.0373	0.0373	0.0538**	0.0544**
		(0.0256)	(0.0264)	(0.0263)	(0.0263)	(0.0255)
	Mean BMI	18.4871	18.4871	18.4871	18.4871	18.4871
		(0.0691)	(0.0691)	(0.0691)	(0.0691)	(0.0691)
	Mean overweight	0.2150	0.2150	0.2150	0.2150	0.2150
		(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)
	N	2,447	2,447	2,447	2,447	2,447
	Child controls		✓	✓	✓	✓
	Mother controls			✓	✓	✓
	Family controls				✓	✓
	Parental BMI					✓

Note: * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003-2006.

Table 2.A.3: OLS estimates: Different definitions of maternal employment

		Dependent		
Employment definition	Mean emp.	Overweight	BMI	N
A: Microcensus				
Not employed	0.3420	*reference*	*reference*	45,210
	(0.0022)			
Part-time	0.4878	-0.0066	-0.0276	
	(0.0024)	(0.0045)	(0.0356)	
Full-time	0.1702	0.0312***	0.3598***	
	(0.0018)	(0.0063)	(0.0509)	
M'whours ≤ 10	0.4685	*reference*	*reference*	45,210
	(0.0023)			
M' whours 10–20	0.2417	-0.0080*	0.0006	
	(0.0020)	(0.0048)	(0.0373)	
M' whours 20-30	0.1210	0.0110**	0.1627***	
	(0.0015)	(0.0063)	(0.0483)	
M' whours ≥ 30	0.1687	0.0340***	0.4008***	
	(0.0018)	(0.0061)	(0.0490)	
M' whours/10	2.3031	0.0140***	0.1457***	30,037
($\mid Employed$)	(0.0074)	(0.0021)	(0.0166)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; M' whours indicates maternal working hours; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009.

Table 2.A.4: Quantile regression estimates: Maternal employment and child BMI

	OLS	Q(0.10)	Q(0.25)	Q(0.50)	Q(0.75)	Q(0.85)	Q(0.90)	Q(0.95)	Q(0.97)
A: Microcensus									
Not employed	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*
Part-time	-0.0276	-0.0117	0.0353	0.0187	-0.0486	-0.0748	-0.0078	-0.0477	-0.1681
	(0.0356)	(0.0394)	(0.0359)	(0.0390)	(0.0489)	(0.0610)	(0.0754)	(0.1123)	(0.1413)
Full-time	0.3598***	0.1694**	0.2984***	0.3740***	0.4148***	0.4388***	0.4074***	0.4050**	0.3587
	(0.0509)	(0.0535)	(0.0487)	(0.0529)	(0.0663)	(0.0827)	(0.1022)	(0.1522)	(0.1916)
N	45,210	45,210	45,210	45,210	45,210	45,210	45,210	45,210	45,210
B: KiGGS									
Not employed	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*
Part-time	0.0100	-0.1788	-0.1368	-0.0317	-0.0552	-0.0158	-0.1930	-0.5855	-0.4721
	(0.1426)	(0.1135)	(0.0949)	(0.1198)	(0.2226)	(0.3638)	(0.3757)	(0.3424)	(0.5556)
Full-time	0.5345**	0.2240	0.3070	0.6114***	0.4779	0.7730	0.7370	1.1493*	0.7651
	(0.1893)	(0.1524)	(0.1273)	(0.1605)	(0.2938)	(0.4669)	(0.4708)	(0.4165)	(0.6375)
N	2,447	2,447	2,447	2,447	2,447	2,447	2,447	2,447	2,447

Note: All models include a full range of child, mother and family controls as well as parental BMI; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

Table 2.A.5: IV estimates: Alternative estimation strategies

Dependent	Mean	(1) OLS	(2) 2SLS	(3) GMM	(4) Treatreg(2st)	(5) Treatreg(ML)	(6) 2SRI	N
A: Microcensus								
BMI	18.3235	0.3765***	2.8150***	2.8114***	2.8114***	4.3432***	1.6931***	45,210
	(0.0162)	(0.0454)	(0.3234)	(0.3234)	(0.3315)	(0.0857)	(0.2572)	
Overweight	0.2324	0.0353***	0.2815***	0.2815***	0.2790***	0.7115***	0.1746***	45,210
	(0.0020)	(0.0056)	(0.0403)	(0.0403)	(0.0416)	(0.0073)	(0.0325)	
B: KiGGS								
BMI	18.4871	0.5285***	1.9116*	1.9092*	1.7663	4.6750***	1.3135*	2,447
	(0.0691)	(0.1683)	(1.1012)	(1.1061)	(1.1544)	(0.3180)	(0.8461)	
Overweight	0.2150	0.0486**	0.2504**	0.2811**	0.2530*	0.6652***	0.1115	2,447
	(0.0083)	(0.0227)	(0.1322)	(0.1333)	(0.1504)	(0.0310)	(0.1128)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; 2RI refers to a two-stage-residual-inclusion approach; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

Table 2.A.6: Robustness analyses: *(5) alternative Instrument* (see Table 2.2)

Model specification		(1) FS M. FT	(2) RF BMI	(3) RF Overweight	N
A: Microcensus					
(5) Dummies: younger siblings					
	None	*reference*	*reference*	*reference*	45,210
	Aged 3	-0.1378***	-0.4257***	-0.0407***	
		(0.0068)	(0.0653)	(0.0081)	
	Aged 3–6	-0.1132***	-0.3142***	-0.0337***	
		(0.0056)	(0.0540)	(0.0065)	
	Aged 6–8	-0.0990***	-0.3037***	-0.0258***	
		(0.0042)	(0.0385)	(0.0049)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; FS refers to the first stage, RF to the reduced form estimations and M.FT refers to maternal full-time employment; The first stage F-statistic is 412.49; Robust standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1999, 2003, 2005, 2009.

Table 2.A.7: Robustness analyses: Household and father's income

Model specification	Dependent	Mean	(1) OLS	(2) FS	(3) RF	(4) 2SLS	N
A: Microcensus							
(1) Without controlling for household income							
	BMI	18.3280	0.3300***	-0.0678***	-0.2120***	3.1290***	45,210
		(0.0166)	(0.0482)	(0.0025)	(0.0229)	(0.3524)	
	Overweight	0.2332	0.0296***	-0.0678***	-0.0213***	0.3142*	45,210
		(0.0020)	(0.0056)	(0.0025)	(0.0029)	(0.0436)	
(2) Including father's income instead of household income							
	BMI	18.3280	0.2939***	-0.0618***	-0.1970***	3.1877***	42,921
		(0.0166)	(0.0482)	(0.0025)	(0.0238)	(0.4011)	
	Overweight	0.2332	0.0240***	-0.0618***	-0.0200 ***	0.3243***	42,921
		(0.0020)	(0.0059)	(0.0025)	(0.0030)	(0.0496)	
B: KiGGS							
(1) Without controlling for household income							
	BMI	18.4871	0.5248***	-0.0875***	-0.1829*	2.0890*	2,447
		(0.0691)	(0.1690)	(0.0153)	(0.1609)	(1.1828)	
	Overweight	0.2150	0.0489**	-0.0875***	-0.0234*	0.2668*	2,447
		(0.0083)	(0.0226)	(0.0102)	(0.0121)	(0.1398)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; FS refers to the first stage, RF to the reduced form and 2SLS to the second stage estimations; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009; KiGGS 2003–2006.

Table 2.A.8: IV estimates: Different age groups

Model specification	Dependent	Mean	(1) OLS	(2) FS	(3) RF	(4) 2SLS	N
A: Microcensus							
(1) Children aged 4–8 years							
	BMI	16.4045	0.2100***	-0.0491***	-0.1252***	2.5527***	53,594
		(0.0161)	(0.0469)	(0.0025)	(0.0277)	(0.5776)	
	Overweight	0.2358	0.0265***	-0.0491***	-0.0154***	0.3143***	53,594
		(0.0018)	(0.0054)	(0.0025)	(0.0032)	(0.0661)	
(2) Children aged 9–12 years							
	BMI	18.3235	0.3765***	-0.0730***	-0.2056***	2.8150***	45,210
		(0.0162)	(0.0454)	(0.0024)	(0.0228)	(0.3235)	
	Overweight	0.2324	0.0353***	-0.0730***	-0.0206***	0.2815***	45,210
		(0.0020)	(0.0056)	(0.0024)	(0.0029)	(0.0403)	
(3) Children aged 13–17 years							
	BMI	20.8452	0.1947***	-0.0849***	-0.1219***	1.4351***	59,084
		(0.0138)	(0.0324)	(0.0022)	(0.0183)	(0.2175)	
	Overweight	0.1717	0.0091**	-0.0849***	-0.0076***	0.0898***	59,084
		(0.0016)	(0.0038)	(0.0022)	(0.0021)	(0.0247)	

Note: All models include a full range of child, mother and family controls as well as parental BMI; FS refers to the first stage, RF to the reduced form and 2SLS to the second stage estimations; * p<0.10, ** p<0.05, *** p<0.01; Robust standard errors in parentheses; *Source:* German Microcensus 1999, 2003, 2005, 2009.

CHAPTER 3

Revisiting the Relation Between Education and Smoking

Joint work with Hendrik Jürges

3.1 Introduction

During the last decades, a vast number of studies dealing with educational differences in smoking behavior have been published. Regardless of the research discipline, a key finding is that lower educated individuals are more likely to smoke compared to higher educated individuals (e.g. de Walque, 2007; Cutler and Lleras-Muney, 2010). More recently, researchers have tried to isolate the causal effect of education on smoking leading to ambiguous evidence (see Grossman 2015 for an overview). While some studies find a strong protecting effect (Jürges et al., 2011; de Walque, 2007; Kenkel et al., 2006), others find no evidence that education impacts smoking behavior (Clark and Royer, 2013; Park and Kang, 2008; Kemptner et al., 2011; Lundborg, 2013). In order to tackle the endogeneity of education most of the studies exploit exogenous sources of variation, for instance changes in compulsory schooling (e.g. Jürges et al., 2011; Clark and Royer, 2013; Li and Powdthavee, 2015) or avoidance of the Vietnam War draft due to college enrollment (Grimard and Parent, 2007; de Walque, 2010), to assess the causal effect on smoking. Although this empirical approach is arguably convincing in identifying a causal effect, its interpretation is confined to a narrow subpopulation of so-called compliers, i.e. those individuals affected by the specific (policy) changes in schooling. The fact that different policy changes affect different subgroups of individuals might to some extent account for the mixed results obtained in previous studies. While these analyses are important in evaluating to what extent expanding education might be successful in preventing the individuals concerned from smoking, one might also be interested in the effect for a broader population.

Moreover, in-depth research is necessary to understand the underlying pathways. If we yield further insights into the mechanisms why smoking is related to schooling, educational policy could be successful in preventing individuals from smoking. Basically, the literature discusses four different — although not mutually exclusive — approaches explaining why higher educated individuals are less likely to smoke or are in better health in general (c.f. de Walque, 2007; Grimard and Parent, 2007; Jürges et al., 2011). First, education might improve the efficiency to produce health inputs, as it is implemented in the Grossman model (Grossman, 1972; Grossman, 2006). A second but complementary explanation is the allocative efficiency theory (Kenkel, 1991; Rosenzweig and Schultz, 1981). According to that theory, education changes the inputs into the health production itself. For instance, higher-educated individuals are more likely to be aware of the harmful effects of smoking or have a higher ability to process health information, such as following medical instructions. Third, education might be viewed as an investment in the future according to human capital theory (Becker and Mulligan, 1997; Becker, 2007). In the context of this explanation, education increases income which makes it more profitable for higher educated individuals to live longer and thus to invest in their (future) health. Finally, apart from these views assuming that education improves health (behavior) through the suggested channels, it might also be that individuals of certain characteristics, like ability or time preference, select themselves both in a higher education and a healthier behavior (Fuchs, 1982).

The classic example supporting the theory that health-related knowledge might prevent individuals from smoking is based on the publication of the US Surgeon's General Report in 1964. Smoking prevalences in the US were declining since this report publicized the harmful consequences of tobacco consumption (de Walque, 2010; Kenkel and Sindelar, 2011). Apart from this historical development indicating that health knowledge might be relevant for smoking decisions, most of the studies that empirically investigate the causal link between education and smoking are confined to a theoretical discussion on this possible mechanism (see Johnston et al., 2015; Pampel et al., 2015). Empirical evidence on health knowledge and other possible mechanisms is scarce. Recent literature emphasizes the importance of peer effects, family background and personality traits for smoking initiation which might also account for the educational differences (Kenkel and Sindelar, 2011; Maralani, 2014; Hsieh and van Kippersluis, 2015).

In this chapter, we investigate three related research questions in order to improve our understanding of educational disparities in smoking. The first and main research question is whether formal education lowers an individual's probability to smoke in a causal sense. Unlike previous economic studies we do not rely on an econometric approach that isolates the causal effect for a complier-specific subpopulation. Instead we apply a descriptive approach with analyses at the population level and examine cohort- and gender-specific educational differences in the probability to initiate smoking for Germany.[28] The analyses rest upon the idea that if formal education affected smoking behavior, educational differences should surface after formal education is completed as the cause (education) must precede the outcome (smoking). We therefore investigate *when* educational disparities in smoking manifest. Making use of retrospective information on the age at smoking initiation we examine age-specific hazard rates to take up smoking. We particularly focus on educational differences in smoking initiation as individuals usually start smoking in early adolescence. We follow the argumentation of Farrell and Fuchs (1982): If educational differences are already apparent at an early age, and thus before education might be effective, the occurring relationship is more likely driven by selection rather than causation. Two studies are similar to our approach in the sense that they focus on educational disparities in smoking initiation age by gender and cohort at the population level. Based on a large sample of the US National Health Interview Survey, Maralani (2013) shows that educational differences in adult smoking trace back to differences in

[28] De Walque (2010) argues that focusing on smoking initiation at early ages misses the dynamics of educational differences in cessation. However, previous literature has shown that differences in smoking are largely attributable to differences in smoking initiation (Jürges et al., 2011; Grimard and Parent, 2007; Maralani, 2013).

smoking initiation in adolescence. Another study by Pampel et al. (2015) examines how educational disparities in smoking vary across cohorts, gender and nations by comparing smoking uptake histories across France, Germany and the US. Such analyses based on German data are of particular interest given the special features of the German selective school system. In Germany, the transition to the secondary school track involves a pivotal career choice. All children attend primary school at least until the age of 10. At that age children change to one of three different tracks of German secondary schools, distinguishing one academic and two vocational tracks.[29] To the best of our knowledge Pampel et al. (2015) is the only study exploiting information on smoking initiation age for Germany. However, this study is based on a relatively small sample of survey data, the German Epidemiological Survey of Substance Abuse. In order to increase precision, we draw on five pooled cross-sections of the German Microcensus, a large representative data source covering more than 1,000,000 individuals. In accordance with the bulk of previous literature, we find strong associations between education and smoking behavior. However, our main analyses suggest that educational differences in smoking initiation are already apparent while individuals are in school. About 85 % (93 %) of the educational differences in smoking among men (women) are determined before the age of 16. Whether individuals ever smoke is predominantly determined at an age before compulsory education is completed. This finding is incompatible with the often assumed strong causal effect of education on smoking. Our results rather support the theory that characteristics determining both the selection into smoking and education, are crucial in explaining educational differences in smoking (Farrell and Fuchs, 1982; Maralani, 2013).

Our second research question relates to these selection mechanisms. Specifically, we investigate whether mechanisms working prior to and during school drive the emergence of educational disparities in smoking. The German Microcensus data lack detailed information on the individual background. We thus draw on another data source, the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) that include a range of observables that might be relevant for this selection. We focus on family resources, peer smoking and the child's well-being. The descriptive analyses reveal that these mechanisms largely account for educational differences in smoking before compulsory schooling is completed.

Third, we explore whether health knowledge acquired as part of a specific health education is important for smoking decisions. For these analyses, we again draw on the German Microcensus data. Whether and to what extent a health education is taught at

[29] The educational system varies slightly throughout Germany as the federal states are responsible for the education policy. The different education systems are still very similar in particular regarding the age at transition to the secondary school track. See Kemptner et al. (2011) for a detailed description of the German school system.

schools remains unclear.[30] We thus rely on post-schooling health education individuals receive as part of their occupational training and studies. Specifically, we compare smoking cessation and initiation rates between health-related occupations, such as physicians, and other occupations with equal educational attainment. The results generally indicate that health education, if at all, has small impacts on the individual's decision to quit smoking. In contrast, individuals working in the health sector were already less likely to take up smoking in adolescence. This finding lends further support that selection rather than causation (running through health knowledge) largely accounts for differences in smoking behavior.

The remainder of the chapter is structured as follows: Section 3.2 briefly summarizes the previous literature and the ambiguous evidence regarding the causal interpretation of the relationship obtained in previous studies. Section 3.3 introduces data and measures. In Section 3.4 we discuss the empirical strategy while Section 3.5 shows the results of these analyses. In Section 3.6 we explore potential mechanisms operating during school while Section 3.7 investigates the importance of a specific health education. Section 3.8 concludes.

3.2 Background

Estimating the relationship between education and smoking behavior, or health in general, is problematic because education is likely to be endogenous (Conti et al., 2010; Maralani, 2014). Aside from the causal effect that higher educated individuals are less likely to smoke, which might be explained by one of the theoretical pathways discussed above, the relationship could also be reversed, reciprocal or evoked by unobserved third factors related to both education and smoking behavior. For example, there might be unobserved individual characteristics, such as time preferences, family background or cognitive ability, that determine education decisions on the one hand and the decision to start smoking on the other hand. Disregarding this endogeneity might in turn result in overestimated correlations between education and smoking. A number of studies try to disentangle the causal effect of education on smoking by exploiting exogenous variation in schooling. However, the evidence obtained by these studies is ambiguous, i.a. depending on (the age of) the subpopulation affected by the exogenous variation in education. For example, de Walque (2007) and Grimard and Parent (2007) find that individuals in the

[30]The Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany publishes recommendations on health education at schools, also in respect of addiction prevention. There exist three different circulations of these recommendations: "Gesundheitserziehung und Schule" published 01.06.1979, "Sucht- und Drogenprävention" 03.07.1990 and "Empfehlung zur Gesundheitsförderung und Prävention in der Schule" published 15.11.2012. See `https://www.kmk.org` for details. But instead of providing specific guidance these recommendations are worded in general terms by primarily describing core competencies the students should posses, similar to the school subject curricula.

US who enrolled for college as a strategy to avoid the draft during the Vietnam War are less likely to smoke. A recent study by Li and Powdthavee (2015) makes use of variations of the compulsory schooling age in Australia and finds that one additional year of schooling does not necessarily prevent individuals from smoking. In their comprehensive study Clark and Royer (2013) exploit two changes in the compulsory schooling in the UK within a regression discontinuity design. Their findings also suggest negligible impacts of education on health and smoking. Based on a comparable instrument for Germany, Kemptner et al. (2011) come to a similar conclusion that there is little evidence of a causal effect of one additional year of schooling for those individuals affected by the change. In contrast, based on the same data, Jürges et al. (2011) find evidence of a strong protective effect of education for individuals that benefited from the educational expansion in the 1950s measured by regional variation in additional grammar school openings in Germany. However, the authors acknowledge that the expansion of grammar schools intended to attract more students, which might have changed the composition of peers with respect to socioeconomic characteristics. For that reason this identified effect of education on smoking might also be driven by peer effects and explain the different results.

Another body of literature focuses on the historical development of educational disparities in smoking prevalences across gender and cohorts (e.g. Pampel, 2005; Piontek et al., 2010; Vedøy, 2014; Bricard et al., 2015). Interestingly, these analyses reveal that the relationship has inversed over time: while smoking was more prevalent among higher educated individuals in older birth cohorts, smoking has become more frequent among the less educated for individuals born 1930–1940 and has been strengthening thenceforward (Schulze and Mons, 2006; Piontek et al., 2010). This pattern can be found for many industrialized countries and is explained by a theory of diffusion which implies mainly three temporal stages: Initially, the adoption of smoking is largely confined to higher educated individuals. A diffusion of smoking to lower educated individuals follows in a second stage while the third stage is characterized by the fact that higher educated individuals begin to reject smoking (e.g. Pampel et al., 2015). Based on this historical development de Walque argues that health knowledge likely accounts for the observed relationship between education and smoking. Analyzing smoking histories from 1940–2000 in the US, de Walque (2010) shows that smoking prevalences generally declined since the harmful consequences of tobacco consumption have become apparent in the beginning of the 1960s. However, this decline started earlier and has been more striking among higher educated individuals leading to the inversion of the educational gradient. De Walque concludes that education facilitates the access to health-related information and/or increases the ability to process this information, according to the productive efficiency theory.

While many studies finding a causal effect of education on smoking theoretically conclude that health knowledge is most likely the predominant pathway, studies that empirically explore *why* education is related to smoking are scarce. In contrast to the conclusion drawn in the study by de Walque (2010), the evidence is less definite when a direct measure of health knowledge is used. Kenkel (1991) finds that individuals with good knowledge about the health consequences of smoking generally smoke less, but this only explains a small part of the observed relationship between general education and smoking. Li and Powdthavee (2015) find no causal evidence that one additional year of education increases an individual's probability to engage in preventive health checks, a behavior which might likely result from some kind of health knowledge. In a recent study based on UK data, Johnston et al. (2015) construct an index of health knowledge comparing the respondent's assessment towards the main causes of ten common health conditions to the answers given by medical professionals. While their OLS model suggest that education is significantly related to better health knowledge, IV estimates based on changes in compulsory schooling indicate that there is unlikely a causal effect of education. Obtaining health-related knowledge might be regarded as a process as it likely takes time and should therefore be rather seen as a mechanism operating later in life.

There is increasing evidence that educational differences in smoking are mostly explained by differences in smoking initiation. Maralani (2013) analyses the role of never smoking to explain cohort-specific educational differences in adult smoking. Based on a large sample of the US National Health Interview Survey, the results show that educational differences in adult smoking trace back to differences in smoking initiation in adolescence. Another study by Pampel et al. (2015) examines how educational disparities in smoking vary across cohorts, gender and nations by comparing smoking uptake histories until age 34 across France, Germany and the US. Their findings generally indicate that results educational differences in smoking uptake are strengthening across cohorts, countries and for both men and women. This recent evidence suggests that individuals typically take up smoking while in school. This observations casts doubt on the prevailing theories that mainly focus on mechanisms operating in adulthood (Maralani, 2013). Only few studies have explicitly investigated early-life characteristics to explain smoking initiation in adolescence. The results suggest that especially social skills and peer effects account for the relationship between education and smoking (Conti and Heckman, 2010; Maralani, 2014; Andersson and Maralani, 2015; Jensen and Lleras-Muney, 2012).

3.3 Data and Measures

In order to investigate educational differences in smoking initiation we use five cross-sections of the German Microcensus 1989, 1999, 2003, 2005, and 2009. The Microcensus is an annual official survey on the living situation of 1 % of the German households covering approximately 800,000 individuals.[31] Although individuals cannot be identified across waves, pooling the five cross-sections leads to a considerably larger sample size which is crucial to have enough observations in particular for early and more recent birth cohorts. Whereas the participation in the Microcensus is mandatory, answering health-related questions, which are usually included every four years, is voluntary. Moreover, before 2005 these health-related questions are asked of a randomly drawn subsample of 50 % in 1989 and 45 % in 1999 and 2003.[32]

For the analyses, we restrict the sample to individuals born between 1930–1989 in West Germany. The birth cohort restriction at the lower bound ensures that there are enough observations in the upper educational group, especially among women. We exclude individuals born after 1989 as they are less likely to have completed their formal education given that the latest survey is from 2009. We only consider individuals living in West Germany to ascertain that the observed differences are not driven by the different educational systems existing in West and East Germany.[33] Finally, we delete observations with non-valid information on any of the variables considered. The analysis sample size amounts to more than 1,000,000 individuals. Sample statistics of the relevant variables used for the whole study population, as well as for the subpopulations of never and ever smoking individuals are presented in Table 3.1.

Measuring Smoking Behavior

The main outcome is smoking uptake. We construct a binary variable that equals 1 for respondents who ever smoked (and thus ever started to smoke) and 0 for never smoking respondents. Within the sample, 48 % of individuals ever smoked. To assess whether educational differences are already apparent at smoking initiation, we make use of retrospective information on the age at smoking initiation answered by current and former smokers.[34]

Measuring Formal Education

We measure formal education as the highest school leaving certificate completed. We generate a dummy variable that equals 1 for individuals who have at least acquired

[31]This official data was provided by the Research Data Centers of the Federal Statistical Office and the Statistical Offices of the Länder in Düsseldorf, Germany, analyzed on-site (further information: http://www.forschungsdatenzentrum.de/en/)

[32]This change in the sampling method also explains the deviation from the four-years cycle of the health information's collection in 2003 and 2005.

[33]However, analyses based on the East German population are similar and lead to the same conclusions (see 3.A.5 in the Appendix).

[34]One limitation of taking retrospective information is that it is prone to response bias as the respondent may not remember the exact age of smoking initiation. Studies have found that in comparison with longitudinal records, recalled information on smoking status were largely accurate (Krall et al., 1989). There is even some evidence that individuals tend to assess the age at onset to be higher than it really was (Bright and Soulakova, 2014).

Table 3.1: Sample statistics (Microcensus), mean values

Variable	All	Never smoker	Ever smoker
Demographics			
Male	0.492	0.406	0.584
Age	45.590	46.346	44.772
Cohort			
1930-1954	0.471	0.423	0.448
1955-1967	0.253	0.327	0.288
1968-1989	0.276	0.250	0.264
Education			
High education: $\geq$ (Fach-) Abitur	0.243	0.275	0.208
High education: $>$ Q.75	0.248	0.286	0.207
Occupation (| same educational level)[a]			
Physicians/pharmacists (academic)	0.060	0.064	0.055
Health-related (intermediate)	0.073	0.075	0.071
%		51.95	48.05
N	1,036,321	538,324	497,997

Note: [a]The sample sizes for the occupation variables differ as it is restricted to individuals with equal educational attainment: 129,311 academic; intermediate 261,413; Column 1 presents the mean values for the whole study population, Column 2 for the subpopulation of individuals who never smoked and Column 3 for the subpopulation of individuals who ever smoked (i.e. currently smoking or stopped smoking); * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

a university entrance qualification (German: Fachabitur and Abitur), and 0 for lower educated individuals.[35] Obtaining a university entrance qualification in Germany usually requires 12 to 13 years of schooling, i.e. about 2 or 3 years supplementary to compulsory schooling. In our data, about 24 % of the individuals are defined as being high educated. The share is higher (28 %) among never smokers.

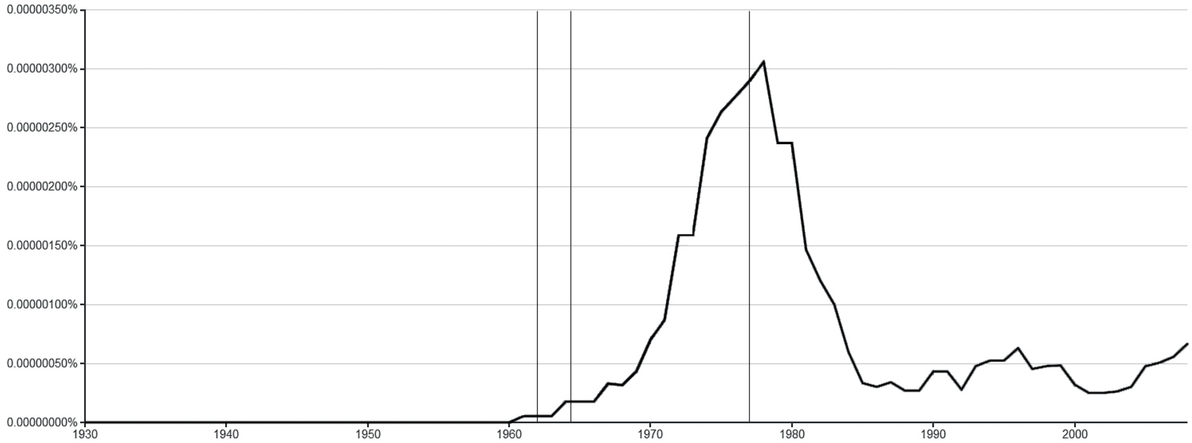
Measuring Cohorts

To assess the development of educational differences in smoking across cohorts we aggregate birth cohorts to three groups. These groups are defined according to the relevant historical events unveiling the harmful consequences of smoking to the public. In order to asses when the harmful consequences of smoking might have become prominent in Germany, we rely on the development of the number of books including the German term "Rauchen und Gesundheit" (engl.: smoking and health) which have been published between 1930 and 2010 (see Figure 3.1).[36]

[35]We use the terms post-compulsory education vs. compulsory education and higher vs. lower education interchangeably.

[36]Unlike the US, there is neither data available to explore smoking prevalences during the last century in Germany nor are we aware of a report analogous to the US report published in Germany.

Figure 3.1: German books published including the term "Rauchen und Gesundheit"



Note: Relative to all books published in a given year; 1962: Royal College of Physicians Report (UK) | 1964: Surgeon's General Report (US) | 1977: ban on tobacco advertising in TV (Germany); *Source:* Google ngram 2016.

Beginning in the early 1960s, the number of books addressing the relationship between smoking and health peaks in 1978 and then declines. This development suggests that also in Germany the public debate on the harmful effects of tobacco consumption was initiated by the reports from the UK (1962) and the US (1964). In 1964 the German news magazine *DER SPIEGEL* took up the discussion and published a special issue on smoking (Der Spiegel, 20.1.1964). This further indicates that these information reached the broader German public in the early 1960s. The debate finally resulted in a policy action in 1977 when a ban on tobacco advertising in German TV and broadcasting has been implemented. For the classification of our cohorts, we assume two historical events to be relevant for the decision to take up smoking for individuals aged 10 and older. The first cohort consists of individuals born between 1930 and 1954, most of whom were unlikely aware of the health-damaging consequences of smoking as the debate was mainly confined to the medical literature. The second cohort (born 1955–1967) comprises individuals whose decision to take up smoking has been made after this knowledge reached the broader public with the publication of the first US Surgeon's General Report in 1964. This report characterizes the beginning of awareness campaigns in the US (de Walque, 2010; Kenkel and Sindelar, 2011) and the debates thus most likely also reached Germany, as the issue of the news magazine *DER SPIEGEL* suggests. Individuals born during this time therefore might have known about the harmful effects of tobacco consumption. Finally, the passing of the ban on tobacco advertising in German TV and broadcast in 1977 marked a time in which the German population as a whole should have been aware of the harmful consequences of smoking. For that reason, the third cohort consists of individuals born between 1968 and 1989, whose decision to take up smoking was likely made in a time when the harms of smoking were known.

3.4 Empirical Approach

We take several steps to explore the main research question, whether formal education lowers an individual's probability to smoke in a causal sense. In a first step, we replicate previous findings and regress ever smoking status on our educational indicator variable applying separate linear probability models (LPM) by birth cohort and gender. That way we explore whether there are significant differences in smoking behavior by education within the data used and how these aggregate differences have changed across birth cohorts.

Second, we explore when educational differences in smoking manifest by exploiting the chronology of treatment (formal education) and outcome (smoking). Although the data used are cross-sectional, we are able to track a part of an individual's life course. The data include retrospective information on smoking initiation and we approximate the individual's age at graduating from school by the age when individuals typically complete their formal education. That is age 16 for compulsory schooling and age 19 for post-compulsory education when individuals receive their university entrance qualification. We argue that if there was a strong causal effect of formal education on smoking, differences in smoking would unlikely occur before formal education is completed. Individuals usually take up smoking in early adolescence. Focusing on educational differences in smoking initiation enables us to estimate whether educational differences are already apparent before the age of 16, i.e. before individuals obtain post-compulsory education. Making use of the retrospective information on the age at initiation, we estimate the education-specific probability that an individual becomes a smoker at a certain age via a discrete time event history model. We thus take smoking initiation as failure event. Specifically, we estimate hazard rates to take up smoking at a given age t:

$$h(t) = \lim_{\triangle t \to 0} \frac{P(t \leq T < t + \triangle t \mid T \geq t)}{\triangle t} = \frac{f(t)}{S(t)}, \tag{3.1}$$

where T indicates the time of event occurrence (which is the age in this context), $f(t)$ is the probability density function and $S(t)$ indicates the survival function, which gives the probability that an individuals has not started to smoke by age t. The width of the age interval (one year) is given by $\triangle t$. We estimate these hazard rates separately by cohort, sex and education. Hence, the hazard rates can be interpreted as the probability that individuals in a given educational group, cohort and sex take up smoking at age t conditional that they have not started before that age. In order to illustrate the different hazard rates for each age graphically, we use the logarithm to better visualize the differences between the two educational groups. The probability to start smoking increases sharply up to the age of 20. While 73 % of the smokers begin at the age of 18 or earlier, less than 5 % take up smoking after the age of 25 within the data used (see

Figure 3.A.1). Accordingly, we estimate hazard rates for smoking initiation at ages 10 to 25. If education had a causal effect on smoking, we would expect the differences in the education-specific hazard rates to increase with age, i.e. with individuals' exposure to education. Smoking initiation after the age of 25 as well as never smoking is right censored. Age 10 is crucial as children usually complete their primary education at this age in Germany. While all children attend primary school, the change to the secondary school track involves a pivotal career choice differentiating into a vocational and an academic track.

In a third step, we make a rough estimate how much of the difference in ever smoking is determined before the age of 16. Until that age all children have to attend school and received quantitatively the same education. Educational differences occurring before the age of 16 can thus likely be attributed to selection rather than causation. This idea is similar to the approach implemented by Farrell and Fuchs (1982). Farrell and Fuchs compare the smoking behavior of individuals before and after they have completed their formal education arguing that a change in behavior would unveil a causal effect of education. Their argumentation can be illustrated by the Rubin causal model (Rubin, 1974; Angrist and Pischke, 2009). According to this theory as presented in Table 3.2, it is impossible to observe the counterfactual smoking outcomes (cells highlighted in gray). For instance, it is not possible to observe the potential smoking behavior of the higher educated if they were low educated. But we do know the individual's age at smoking onset. We thus approximate the counterfactual smoking outcome of high educated individuals by their smoking behavior before the age of 16, i.e. before the higher educated separate from the lower in terms of three additional years beyond compulsory schooling (cell shaded in light-gray).

Table 3.2: Potential outcomes framework

	Schooling		
	Compulsory	Post-compulsory	*Diff.*
Low education	$E[Y_i^0 \mid educ = 0, X]$ smoking outcome of low educated until compulsory education is completed	$E[Y_i^1 \mid educ = 0, X]$ smoking outcome of low educated after post-compulsory education is completed	*ATC*
High education	$E[Y_i^0 \mid educ = 1, X]$ smoking outcome of high educated until compulsory education is completed	$E[Y_i^1 \mid educ = 1, X]$ smoking outcome of high educated after post-compulsory education is completed	*ATT*
Difference	*Selection bias*	*Selection bias*	

Note: The gray-shaded expected values highlight the counterfactual outcomes. X indicates a vector of observables, such as the individual's age.

We consider the educational difference in the probability to take up smoking before the age of 16 (compulsory schooling age). Specifically, we estimate how much of the total difference in ever smoking can be explained by smoking differences before higher education is realized. This difference can be interpreted as selection bias. We define the total difference in ever smoking at the age of 25 as individuals have most likely finished their schooling and only few individuals take up smoking after that age (see Figure 3.A.1). It has to be considered that the approximation of the expected smoking outcome are also conditional on other factors X, such as age. Our approximation of the counterfactual for high-educated individuals (shaded in light-gray) is still appropriate to roughly examine the selection bias at the population level as we are able to keep age constant. However, this assumption does obviously not apply for comparisons between compulsory and post-compulsory educated individuals. That is because we are unable to estimate the age-specific smoking status for each individual as the data lack information on the age at smoking cessation. For that reason our approximation of the counterfactual (shaded in light-gray) is inappropriate to examine the causal effect (ATT) directly. Our approach thus slightly deviates from the approach implemented by Farrell and Fuchs (1982). We are not able to directly assess if higher education induces individuals to change their smoking behavior.

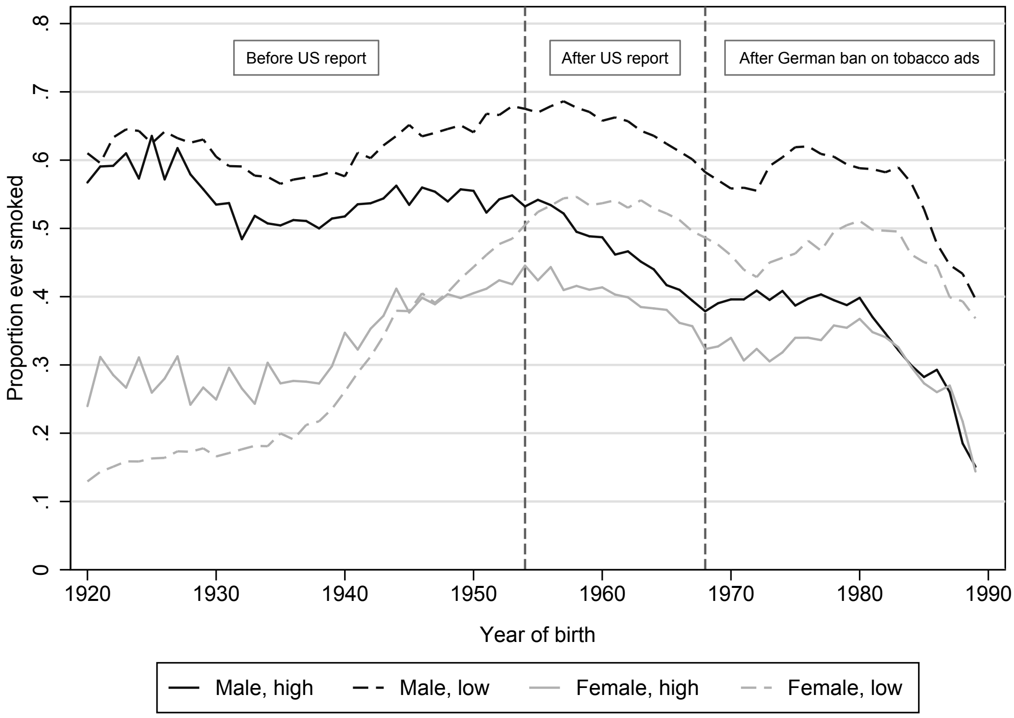
We perform all analyses separately by birth cohort and gender because previous studies have shown that educational disparities in smoking have changed across time and birth cohorts (e.g. de Walque, 2007; Maralani, 2013; Piontek et al., 2010). We do not control for further variables apart from birth cohort and gender. That is because we aim to examine raw educational differences in smoking exploiting retrospective information on smoking initiation. Individuals usually start smoking in adolescence and thus most of the observable variables are likely to be endogenous as they manifest later in life.

3.5 Results

3.5.1 Educational Differences in Ever Smoking Across Cohorts

First, we explore whether significant differences in ever smoking between low and high educated men and women exist and whether cohort changes are apparent within the data used. Figure 3.2 illustrates cohort trends in ever smoking by gender and education. The smoking prevalences increase for individuals whose decision to take up smoking was likely made at a time when the harmful consequences of smoking were largely unknown (before US Report). The share of ever smoker peaks for individuals born in the 1950s and generally declines from then on. The sharp decline in the prevalences for men and women born after 1985 likely arises due to a composition effect. Individuals born in the late 1980s are mainly observed in the more recent waves of the German Microcensus. These individuals are thus relatively young (aged 16–20) compared to earlier cohorts and still might take up smoking, i.e. become an ever smoker, after the end of the study. The general pattern reflects that the publication of the US Surgeon's General Report might have been crucial for Germany as well. While earlier-born men tend to ever smoke more often than women, these gender differences are narrowing within each educational group across cohorts. With respect to the educational differences in ever smoking, Figure 3.2 illustrates a changing pattern across birth cohorts: While there are hardly any differences by education in ever smoking among men from older birth cohorts, prevalences are diverging across cohorts with higher smoking rates among low-educated men. For women, smoking has been more prevalent among higher educated women born up to 1945 but educational differences have inverted from then on. The disparities appear to be most pronounced for individuals born in the most recent year considered, with slightly greater differences among men than women.

Figure 3.2: Cohort trends in ever smoking by gender and education: Prevalences



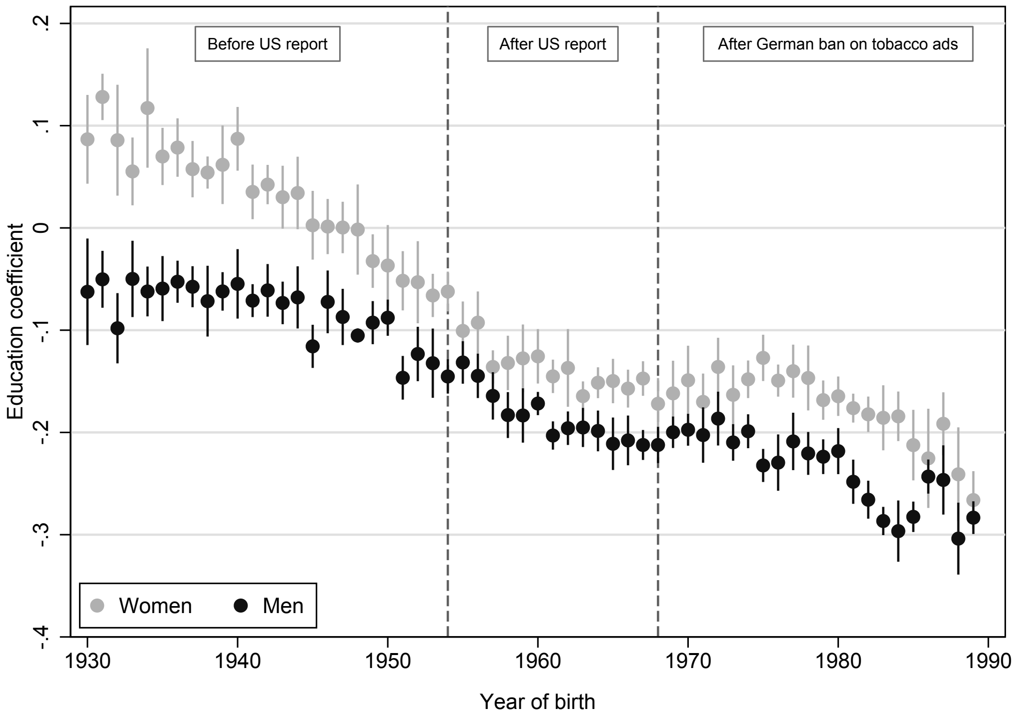
Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

The cohort pattern persists when we estimate educational differences in ever smoking via a simple LPM model controlling for German nationality, fixed effects for states of residence and annual birth cohorts as well as a firth order polynomial in age. Figure 3.3 plots the coefficients and the corresponding 95 % confidence intervals for high education obtained from separate regressions by year-of-birth and gender.[37] As expected from the illustrations in Figure 3.2, educational differences have increased across birth cohorts and

[37] The OLS results for the aggregate measure differentiating the three different cohorts are presented in Table 3.A.1 in the Appendix.

are more severe for men than for women. While high educated men born in 1930 have a 6 percentage point lower probability to ever smoke, their later-born counterparts from the 1980s have a 25 percentage point lower probability. The educational differences did not change substantially for about fifteen years after the health hazards of smoking became publicly known. But for individuals born after 1975 the educational differences further increase. The development is similar for women, although high educated women born between 1930–1945 were even more likely to ever smoke compared to lower educated women of this generation. The observation that the estimates become more negative across birth cohorts — indicating that the younger the birth cohort the stronger the relationship between education and ever smoking — is in line with previous findings from other countries (e.g. de Walque, 2010).

Figure 3.3: Cohort trends in ever smoking by gender and education: OLS estimates



Note: Each circle/diamond presents an education coefficient and the corresponding 95 % confidence interval obtained from seperate OLS regressions of ever smoking on high education; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parantheses; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

3.5.2 Educational Differences in Smoking Initiation

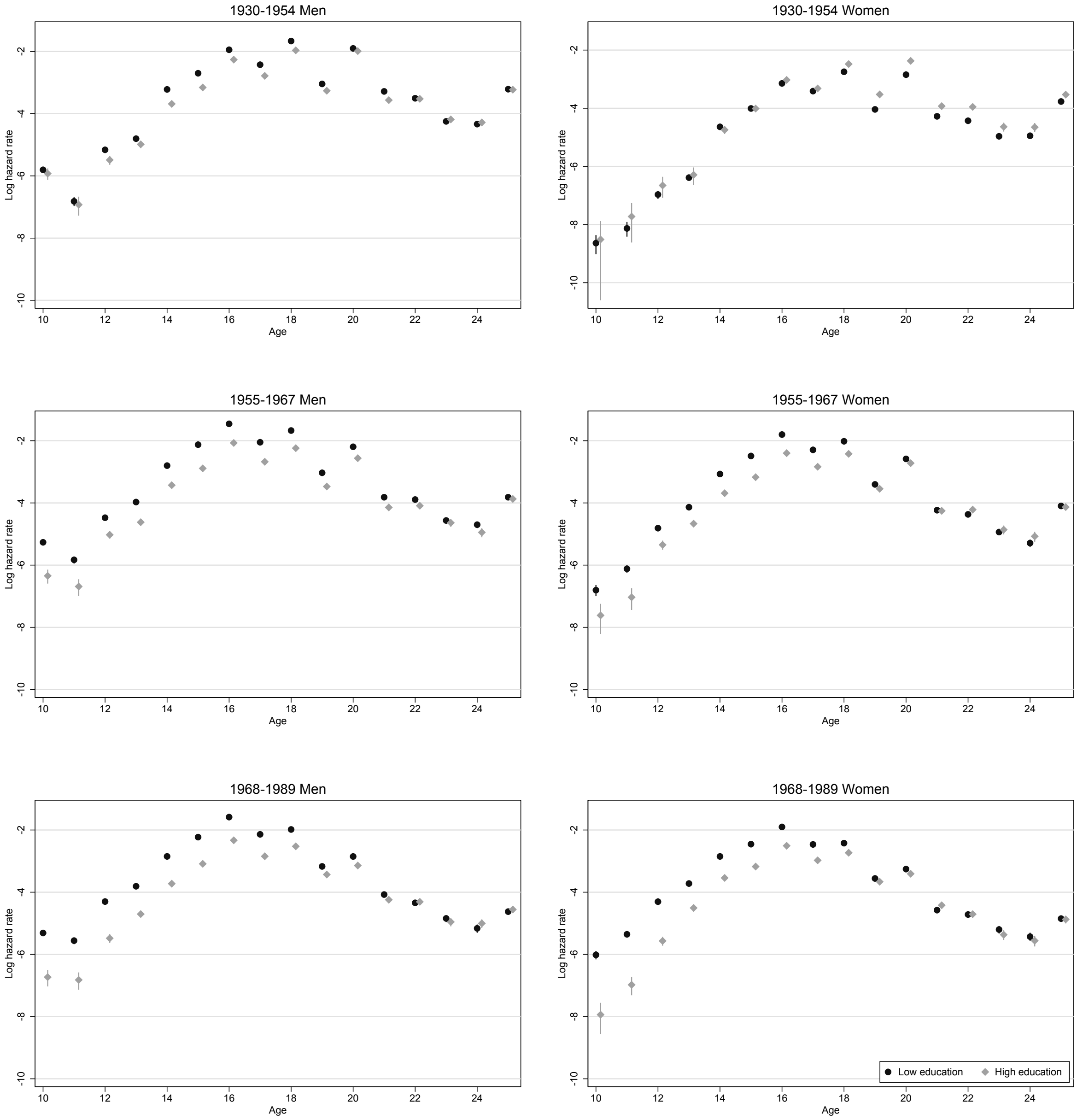
In accordance with the previous literature we find strong associations between education and the probability to ever smoke, which have become more severe across cohorts. However, considering the ever smoking prevalence is ambiguous as it remains unclear *when* the educational differences in smoking manifest. For this reason we now focus on educational differences in smoking initiation. Figure 3.4 illustrates the estimated (log) hazards of taking up smoking at age 10 to 25 and the corresponding 95 % confidence intervals separately for low educated (black circles) and high educated men and women (gray diamonds) and the three cohorts. In general, the hazard rates follow a similar age pattern: The probability to start smoking increases steadily from age 10, peaks at ages between 16 and 20 and then decreases. Individuals aged 16 to 20 have a log hazard rate around -2 which corresponds to a probability to start smoking of about 15 %. For earlier cohorts, the pattern is slightly shifted to the right. This indicates that older individuals tend to have initiated smoking somewhat later. The obvious peaks at age 16, 18, 20 and 25 might be explained by recall bias (e.g. rounding to even numbers), as the analyses are based on retrospective information. This especially holds for older cohorts, where these peaks are more pronounced. Alternatively, the peaks at age 16 and 18 could also be evoked by the fact that in Germany, 16 is the legal smoking age and 18 the legal age in general. For women the age pattern is slightly shifted to the right as women tend to take up smoking a little later compared to men.

With respect to education, we find significantly higher hazard rates to start smoking at an early age for lower educated compared to higher educated men. For women, the differences across cohorts are more pronounced. While there are hardly any educational differences in the probability to start smoking for women born in 1930–1954, the differences become apparent for later born women. For both, men and women, the educational differences become smaller and insignificant in adulthood. This pattern is similar across cohorts although it becomes more explicit for younger men and women. This observation is in accordance with the widening educational gradient in smoking across cohorts (cf. Figures 3.2 and 3.3). Interestingly, the differences in smoking probabilities diminish over the life course, i.e. the longer an individual has been in school and thus the more education it has acquired.[38] Educational differences in smoking are largest before education is completed and even before the minimum school leaving age of 16 is reached. After age 20, and thus at an age when individuals already have acquired their university entrance qualification (post-compulsory education), the differences between the two educational groups become negligible and statistically insignificant.

In conclusion, our results are in line with the previous studies performing similar analyses (Maralani, 2014; Pampel et al., 2015) and show that differences in smoking by completed education already occur at age 10 and thus way before education is completed.

[38]This pattern becomes more clear-cut when we illustrate the differences in hazard rates (the log hazard rate ratios), see Figure 3.A.3 and 3.A.4.

Figure 3.4: Log hazard rates: Smoking initiation by completed formal education

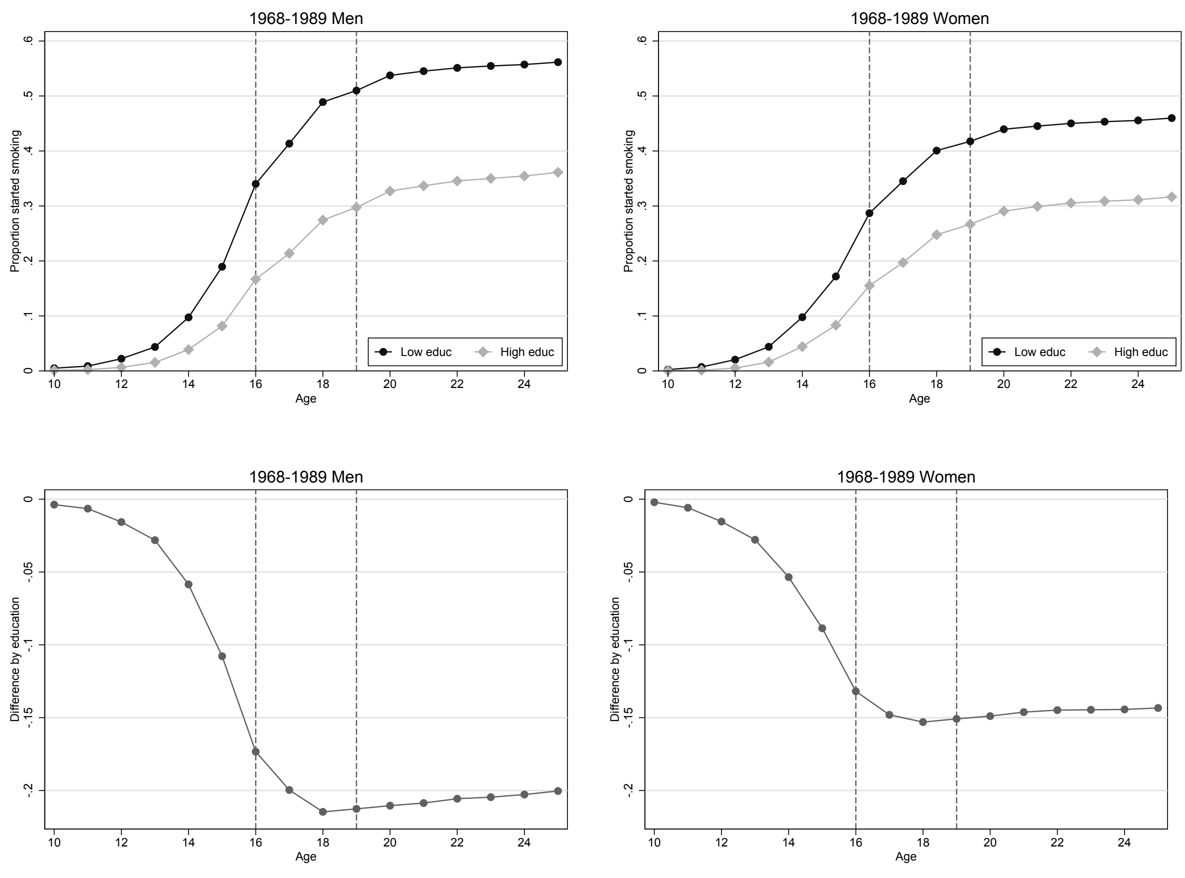


Note: Each circle/diamond presents the age-specific log-hazard rate and the corresponding 90 % confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

3.5.3 Decomposition: Selection vs. Causation

We next try to disentangle the educational differences in smoking initiation before the age of 25. We define the total difference in ever smoking at the age of 25 as individuals have most likely finished their schooling and only few individuals take up smoking after that age (see Figure 3.A.1). As discussed in Section 3.4 we argue that educational differences in smoking up to the age of 16 most likely result from pre-treatment selection characteristics and resulting differences, e.g. in the quality of peers. For convenience, we focus on individuals of the most recent cohort born between 1968 and 1989 as educational differences in smoking are most pronounced for this group.[39] Figure 3.5 shows the cumulative proportion of men and women who took up smoking before a given age by education (upper panel) and its calculated differences (lower panel).[40]

Figure 3.5: Educational differences in smoking initiation until a given age



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

[39]See Figure 3.A.6 and 3.A.7 in the Appendix for the results of earlier cohorts.

[40]Unlike the hazard rates (e.g. Figure 3.4), these proportions have to be interpreted in a cumulative way. The education-specific hazard rates, which estimate the probability to become a smoker at a specific age, indicate the slope of these distributions at a given age (cf. Equation 3.1).

In line with the previous analyses, the proportions in taking up smoking increase with rising age for both educational groups and are again higher for men than women. The share of taking up smoking increases up to the age of 18 but rises more sharply for the low educated individuals. After that age, both curves run almost parallel and level off. Accordingly, the calculated differences between high educated and low educated individuals do hardly change after the age of 18. Although the pattern is similar across men and women, the educational differences are again more pronounced for men. One might be concerned that the figures are driven by the fact that higher educated individuals just take up smoking at later ages. However in this case, the curve of the higher educated would catch up the other curve at a certain age. This is obviously not the case as the curves run parallel after the age of about 18.

At the age of 25 the difference in smoking between high and low educated individuals amounts to 20 percentage points for men and 14 percentage points for women. Considering the educational difference in taking up smoking before the age of 16 reveals that the difference is already on a high level at that age: 17 percentage points for men and 13 percentage points for women. In other words, the educational differences in smoking at the age of 16 constitute 85 % for men and 93 % for women of the total difference in ever smoking at the age of 25. Following the discussion in Section 3.4, we argue that if at all, at most 15 % (7 %) of the differences in ever smoking between high and low educated individuals might be attributed to the *causal* effect of post-compulsory education. Ongoing peer effects might also (partly) be responsible for the remaining difference after the age of 16. Regardless of the factors driving this residue, the decomposition suggests that the major part is likely explained by selection factors and resulting (peer) effects which are determined pre-school, at the age of 10 or even before that age.

Although higher and lower educated individuals likely received quantitatively the same education one might be concerned that the results are driven by disparities in the quality of education. We cannot directly explore whether quality differences might cause the observed educational disparities in smoking. However, the results presented in Section 3.5.2 suggest that the relative differences in smoking initiation rates are largest at age 10 and diminish over the life course, i.e. the longer individuals received education. If there were considerable differences in the quality of education across the two educational groups, we would expect these relative differences to increase over time. As our results indicate the opposite, we conclude that quality differences in education unlikely account for our findings.

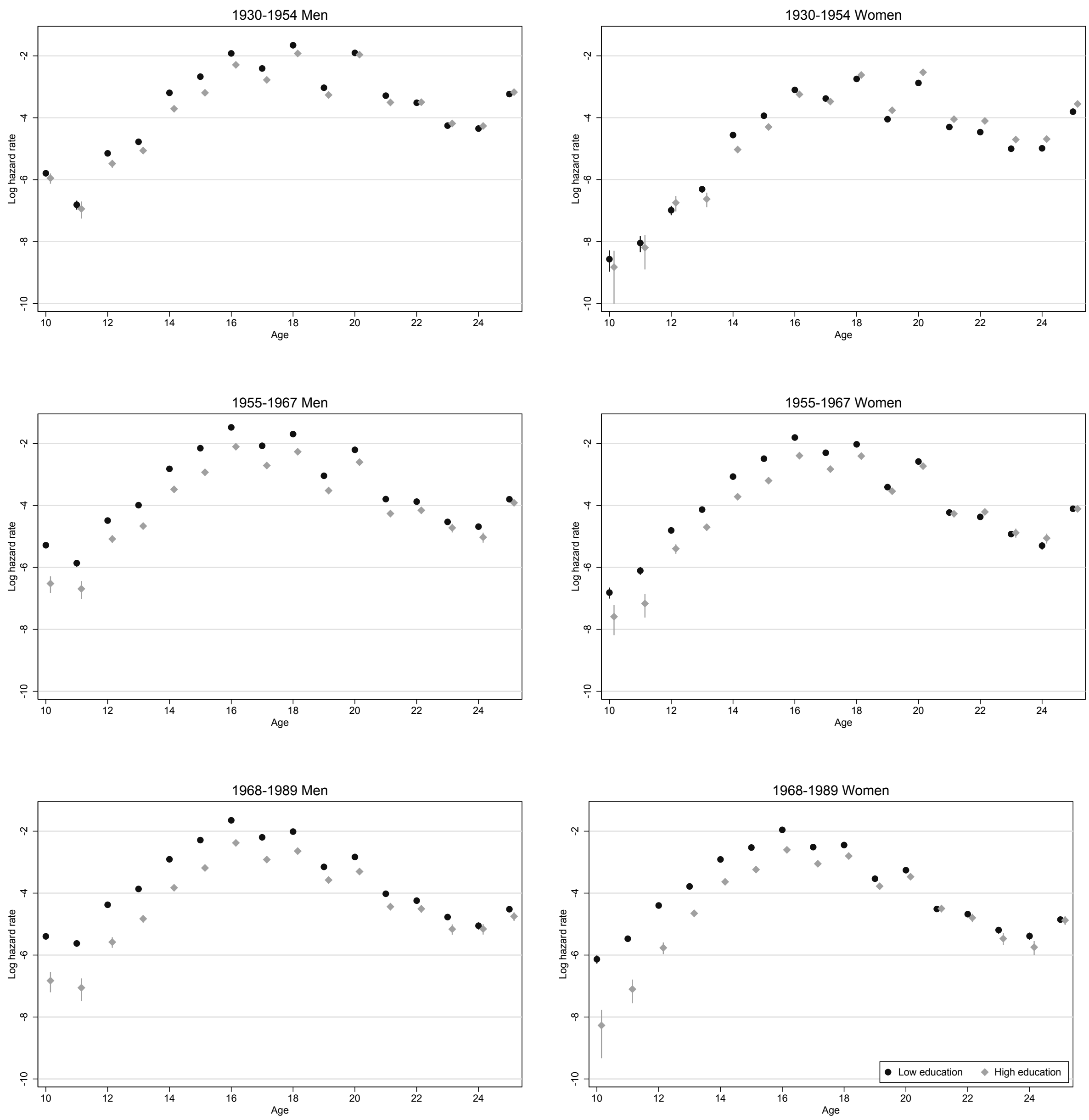
3.5.4 Robustness Analyses

So far, the preferred measure of education is based on the formal secondary school qualification. However, the share of individuals with university entrance qualification has risen at least since the educational expansion in the 1950s in Germany. Due to this improved access to higher education, earlier-born individuals obtaining their university entrance qualification might be more positively selected in terms of individual ability, preferences or socioeconomic background compared to their later-born counterparts (Bound et al., 2014). This may bias our results as those individuals might also be different with respect to their smoking behavior. To address this selectivity we a use a relative measure of education following Jürges et al. (2016) to check the robustness of our results. First, we assign years of education to each individual according to its school and vocational education and generate age- and gender-specific groups of education.[41] Finally, we define individuals as higher educated if they belong to the upper quartile of their age. We thus take the upper 25 % in terms of their years of schooling of each cohort. The advantage of this approach is that we are able to compare these individuals according to their educational ranks which are comparable over time. This definition is comparable to the main specification regarding the share of higher educated individuals. Although we cannot clearly distinguish individuals with compulsory and post-compulsory schooling any longer these analyses are important to assess any potential bias.

Figure 3.6 illustrates the findings for this relative measure of education. The hazard rates follow the age pattern which is already known from the main analyses. While the probabilities to take up smoking increase until the age of 16 to 20, they decline thenceforward. The pattern is again somewhat shifted to the right for earlier-born individuals. Regarding the educational differences, we find small differences for women born between 1930 and 1954. Compared to the main analyses, the hazard rates for high educated women are somewhat smaller. This indicates that within this cohort, the main results might slightly be driven by the improved access to higher education for later-born women. This seems to be plausible as especially those women might have benefited from the educational expansion. For younger cohorts, the hazard rates for high educated men and women are slightly smaller at older ages. However, the overall pattern suggests that the results based on the relative education measure are very similar compared to the main results. We also performed these analyses with a different definition of relative education, defining individuals of the upper half as higher educated (results not shown). The results are again very similar which indicates that the findings are also robust to a shift in the threshold.

[41] Years of schooling are not directly available in the Microcensus data. However, we follow Kemptner et al. (2011) and use the number of years usually spent in school according to the individual's secondary school degree.

Figure 3.6: Robustness: Log hazard rates for smoking initiation by relative education



Note: Each circle/diamond presents the age-specific log-hazard rate and the corresponding 90 % confidence interval obtained from a discrete time event history model taking smoking initiation as failure event; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

3.6 Mechanisms During School (KiGGS Data)

The previous analyses reveal that educational differences in ever smoking are largely determined before formal schooling is completed. It is yet unclear what (selection) mechanisms operating prior to and during secondary education account for these differences. The German Microcensus data lack information on the individuals' childhood conditions. To assess what mechanisms might account for the early determining of ever smoking we draw on the German Health Interview and Examination Survey for Children and Adolescents (KiGGS).

Data and Methods

The KiGGS study is a nationally representative survey on the health of 17,641 children aged 0–17 years and was conducted 2003–2006.[42] Although the KiGGS is also cross-sectional, it includes much more observables on the child's family and individual background. The data enable us to explore potential mechanisms operating prior to and during secondary school attendance. We restrict the analyses to children attending secondary school until grade 10 (i.e. before compulsory schooling is completed) with non-missing data.[43] The analysis sample amounts to 3,794 children.

We estimate age-specific OLS models for the relationship between education and current smoking. We consider current smoking status as dependent variable, that equals 1 if a child currently smokes and 0 otherwise. Our educational measure equals 1 if a child attends an academic track school and 0 for students attending a basic or intermediate track school. Given the early educational selection at the age of 10, we are able to distinguish children who will likely receive post-compulsory schooling before they indeed received this higher education.[44] The education coefficient can thus be interpreted as the difference in the smoking prevalences between children attending academic track schools and children in basic track schools at a given age. Adding three different sets of controls successively enables us to assess whether certain characteristics mediate the relationship between type of school and smoking. We relate to the mediation framework suggested by Baron and Kenny (1986) although we confine ourselves to present the results for the main paths and compare the education coefficients across the different models. We estimate separate models by age to assess when certain characteristics are most relevant while allowing them to vary with age.

[42] See Kurth et al. (2008) for detailed information on the KiGGS data.

[43] We drop 93 students aged 17 who still attend grade 10 as we perform age-specific analyses.

[44] It has to be mentioned that some children might switch between the different secondary school types before compulsory schooling is completed. However, children most often change from the academic to the basic school track. If this at all biases our results, the estimates can be interpreted as lower-bound estimates.

The selection of the potential mechanisms is generally motivated by the previous literature that highlights peers, parents and money to be the key influences (Kenkel, 2012; Jensen and Lleras-Muney, 2012; Maralani, 2014).[45] We distinguish three different sets of covariates which are classified according to their timing of determination. We consider demographics as baseline characteristics, namely the child's sex and residence as well as his or her ethnic background defined by the parents' nationality. In a second set of covariates, we include variables capturing the child's family background that is most likely determined before the age of 10 and thus pre-determined to both, the transition to secondary schools and the child's decision to take up smoking. Specifically, we consider family socioeconomic status and whether or not the mother and father currently smoke. We distinguish between low, medium and high family socioeconomic status based on a multi-dimensional index combining parental education, occupation and income (see Lampert et al. 2014 for details). Although we have no information when the parents started to smoke, our previous analyses suggest that few individuals take up smoking in adulthood. For that reason we interpret parental smoking status to be determined prior to the child's smoking and education decision. We finally summarize subjective characteristics that are significantly related to both smoking and the type of school. Specifically, we include whether a friend smokes or not to capture peer effects on smoking, whether the child is risk-loving (proxied by not wearing a protective helmet while bicycling) and the monthly money at hand, combining self-earned and pocket-money. Moreover, we also include three measures to capture the child's psychological well-being. We include a variable indicating whether or not the child tends to have behavioral problems[46] as well as two summary scales for the child's self-assessed well-being regarding family and friendship.[47] It has to be considered that these variables are likely subject to simultaneity bias and the direction of the causal link is totally unclear.[48] However, the purpose of the analyses based on the KiGGS data is to make a first approximation what mechanisms likely drive the educational differences in smoking in a purely descriptive sense. For that reason we include these potentially endogenous characteristics but do not claim to establish a causal effect on smoking.

Table 3.3 reports the mean values of the variables for all children, separately for children in the basic and academic school track as well as its raw difference (academic –

[45]See Kenkel (2012), Jensen and Lleras-Muney (2012) and Maralani (2014) for a detailed discussion on potential mechanisms operating during school.

[46]The measure of behavioral problems is derived from the "Strength and Difficulties Questionnaire (SDQ)" (Goodman, 1997). The other SDQ-scores (emotional symptoms, hyperactivity/inattention, peer relationship problems and pro-social behavior) were no longer significantly related to smoking nor did the coefficient of the main predictor changed once we controlled for behavioral problems. For that reason we decided to omit the other SDQ-scores.

[47]These measures are subscales from a quality of life instrument for children (Ravens-Sieberer and Bullinger, 1998). We do not consider the other subscales as they are generally unrelated to smoking and thus do not match the requirements of a mediator variable.

[48]See Manski (1993) for a discussion on the problems when estimating peer group effects.

basic). While 19 % of the children aged 11–16 within the basic school track smoke, the prevalence is half as big among children in the academic school track.

Table 3.3: Sample statistics (KiGGS), mean values

		School track		
Variable	All	Basic	Academic	Difference
Smoking				
Current smoker	0.149	0.187	0.092	-0.095***
Demographics (baseline)				
Age 11	0.157	0.151	0.167	0.016
Age 12	0.169	0.157	0.189	0.032**
Age 13	0.186	0.189	0.182	-0.007
Age 14	0.189	0.186	0.193	0.007
Age 15	0.182	0.190	0.169	-0.021
Age 16	0.116	0.127	0.100	-0.028***
West Germany	0.312	0.294	0.339	0.045***
Boy	0.518	0.552	0.467	-0.085***
Mother: German	0.933	0.922	0.949	0.027***
Father: German	0.927	0.915	0.945	0.030***
Family background				
SES: low	0.219	0.313	0.077	-0.236***
SES: medium	0.509	0.540	0.462	-0.078***
SES: high	0.272	0.148	0.462	0.314***
Father smokes	0.382	0.437	0.297	-0.140***
Mother smokes	0.304	0.352	0.232	-0.121***
Potentially endogenous characteristics				
Friend smokes	0.396	0.461	0.297	-0.164***
Risk-loving	0.677	0.711	0.625	-0.086***
Money (in Euro, per month)	31.76	32.92	29.98	-2.943**
Behavioral problems	0.277	0.326	0.202	-0.124***
Well-being: friends [0;100]	82.27	78.44	77.63	-0.810*
Well-being: family [0;100]	78.12	81.78	83.02	1.249**
N	3,794	2,291	1,503	
%		60.38	39.62	

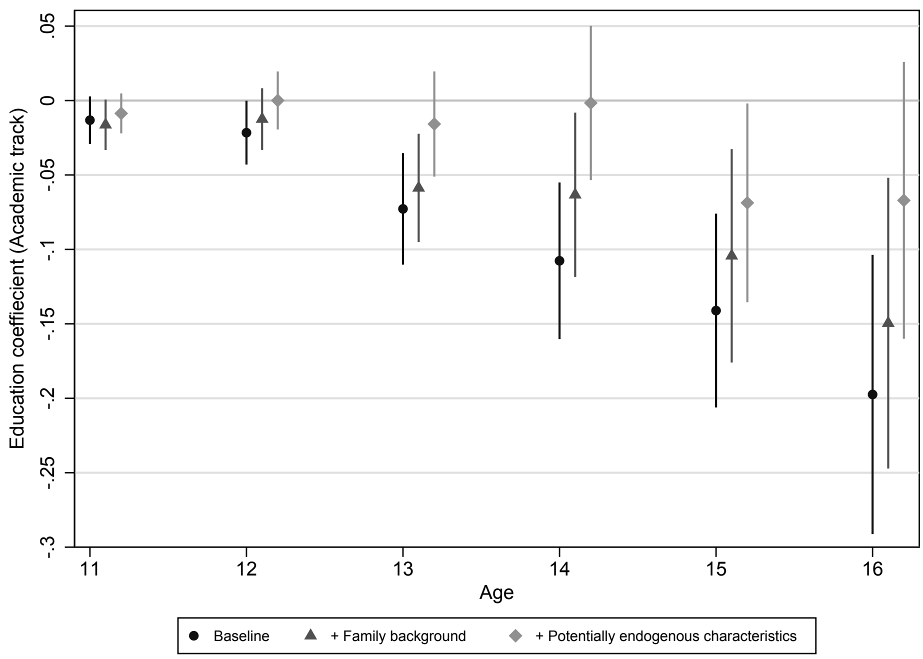
Note: Column 1 presents the mean values for all children, Column 2 for children in basic track schools, Column 3 for children in academic track schools and Column 4 its calculated difference (academic–basic); The variables are dichotomous [0;1] unless specified otherwise; * p<0.10, ** p<0.05, *** p<0.01; *Source:* KiGGS 2003-2006.

The mean values clearly indicate that children attending academic track schools are more likely to have a favorable family background. Whereas 46 % of the children in the academic school track have higher socioeconomic background, this only applies to 15 % of the children in basic track schools. Additionally, children in lower education tend to have smoking parents more often. Regarding the potentially endogenous characteristics children in basic track schools also seem to have adverse conditions: Compared to children seeking a higher educational degree they have a smoking friend more often and are more likely to have behavioral problems.[49]

Results

Although the considered characteristics are unequally distributed across children in the two school tracks it remains unclear if they account for the educational differences in smoking. Figure 3.7 illustrates the coefficients for high education (attending the academic school track) on current smoking obtained from age-specific OLS regressions. We estimate three different models adding the sets of covariates successively. Table 3.4 additionally provides the corresponding adjusted R^2 to evaluate the descriptive importance of the covariates included to predict current smoking.[50] As reference, the table also includes the age- and education-specific smoking prevalences to better evaluate the estimated coefficients.

Figure 3.7: OLS estimates: High education and current smoking (during school)



Note: Each circle (triangle/diamond) presents the coefficient of attending an academic track school on current smoking and its 95 % CI obtained from separate OLS regressions; *Source:* KiGGS 2003-2006.

In general, the OLS results (Figure 3.7) suggest that the educational difference in smoking is small and indistinguishable from zero at an early age but increases substantially until the age of 16. Conditional on demographics (baseline), children attending

[49]These characteristics correlate significantly with current smoking in the following directions: friend smokes (+), risk-loving (+), well-being w.r.t family (-), well-being w.r.t friends (+), tending to have behavioral problems (+), money at hand (+)

[50]See, e.g. Kvalseth (1985) or Anderson-Sprecher (1994) for a critical discussion on R^2.

an academic track school are 18 percentage points less likely to smoke at the age of 16 compared to children in the basic track. This finding is quantitatively comparable to the educational differences in ever smoking obtained from the German Microcensus for the most recent birth cohort (see Table 3.A.1 in the Appendix). The observation that educational differences in smoking increase with rising age is in line with our previous analyses suggesting that the gap in smoking is already widening while individuals attend school. When we consider the smoking prevalences at age 16, about one in two students attending the basic school track smoke. In contrast, only one out of three children smoke among the academic-track students (Table 3.4, upper panel). The estimated difference shrinks after controlling for family background characteristics. However, the coefficients are still significant for children aged 13 and older. The reductions in the estimates itself turn out to be not statistically significant. But when the potential endogenous characteristics are additionally included, the estimates reduce twice as much. Compared to the baseline model the coefficients reduce significantly at the 95 % level for children 12 to 16. Except for age 15, the education coefficients also become indistinguishable from zero. However, this might also result from the large confidence intervals as the coefficients at age 15 or 16 are quantitatively still sizable.

A comparison of the adjusted R^2 (Table 3.4, lower panel) across the three models also point to the importance of the potentially endogenous characteristics. The demographic variables included in the baseline model seem to “explain” little of the variation in current smoking across all ages compared to the null model. When we control for family background, the “explained” proportion doubles on average while it redoubles when the potential endogenous characteristics are included. For instance, while the baseline model accounts for 4 % of the smoking variation at age 16, this proportion increases to 9 % when controlling for family background and to 25 % after including the potentially endogenous characteristics.

It has to be kept in mind that these results should be interpreted with caution. That is because we are unable to distinguish between whether the smoking behavior of the child of interest influences these characteristics, e.g. peer smoking, or the other way around. Taken as a whole, the results suggest that especially mechanisms operating during school are crucial for the individual’s decision to smoke. In analyses including the potentially endogenous characteristics successively (results not shown) it becomes apparent that whether a friend smokes seems to be the most crucial mechanism. This is in line with previous literature from other countries that finds smoking peers or social networks in general to be an important pathway operating in adolescence (Jensen and Lleras-Muney, 2012; Maralani, 2014; Andersson and Maralani, 2015). But in accordance with these studies, we also conclude that the included characteristics do not fully explain smoking prevalences in adolescence, at least not at an older age. There seem to be other unobserved and elusive factors that are particularly important for students attending

Table 3.4: Age-specific smoking prevalences and adjusted R^2 (cf. Figure 3.7)

	Age					
	11	12	13	14	15	16
Basic school track						
Current smoker	0.017	0.028	0.104	0.209	0.308	0.493
Friend smokes	0.119	0.212	0.367	0.592	0.699	0.767
N	346	359	433	426	435	292
Academic school track						
Current smoker	0.004	0.011	0.037	0.110	0.185	0.300
Friend smokes	0.060	0.095	0.197	0.379	0.571	0.640
N	251	284	274	290	254	150
OLS Models: adjusted R^2						
Baseline	0.010	0.009	0.012	0.035	0.033	0.044
+ Family Background	0.018	0.014	0.018	0.075	0.075	0.099
+ Potentially endogenous	0.063	0.071	0.112	0.216	0.211	0.254
N	597	643	707	716	689	442

Note: * p<0.10, ** p<0.05, *** p<0.01; *Source:* KiGGS 2003-2006.

basic track schools. Moreover, health knowledge obtained via a specific health education might also account for the (remaining) differences. We try to examine the importance of a specific adult health education in the next section.

3.7 Post-Schooling Health Education

Finally, we assess whether health knowledge acquired post-schooling as part of a specific health education is relevant for an individual's smoking decision. Again, we base these analyses on the German Microcensus data. The previous analyses suggest that general education itself is unlikely the driving force behind the differences in smoking initiation. Individuals initiate smoking before the formal educational degree is obtained leading to a reversed order of cause and effect. Further, while mechanisms operating during secondary school, such as peer smoking, account for a large part, they fail to fully explain why individuals start to smoke, especially after age 14. A certain health education, as part of formal education might still be relevant for the decision to quit smoking. This decision is more likely made in adulthood, after the educational degree is obtained. Previous studies have shown that among those individuals who initiated smoking, higher educated individuals are also more likely to stop smoking. However, evidence on its causal effect is again mixed (e.g. Jürges et al., 2011; Kemptner et al., 2011; de Walque, 2007). Based on the German Microcensus data we also find that individuals who received post-compulsory education are more likely to stop smoking compared to lower educated individuals (see

Table 3.A.2). Unlike the findings for ever smoking, these associations are stronger for women but remain largely stable across birth cohorts.[51]

We argue that these educational differences in smoking cessation might also result from unobserved factors related to the selection into higher education and smoking cessation (rather than smoking initiation). But given that the decision to quit smoking is likely made after the age of 16, these differences might be indicative for a causal effect of education, mediated by acquired health knowledge. Whether and to what extent a health education is taught at schools remains unclear. It is thus obscure whether the students explicitly learn what lifestyles are harmful or beneficial for an individual's health.[52] But regarding the education individuals receive as part of their studies we do know the teaching contents in a large part. For physicians and pharmacists the content is health-related by definition. For that reason, we rely on post-schooling health education to explore its role for smoking decisions.

Data and Methods

We again draw on the German Microcensus data which is described in Section 3.3. In order to examine whether health-related education is important for smoking behavior we focus on health education obtained post-schooling. We define individuals who acquired specific health education as individuals working in the health-related sector using information on occupations according to the German classification of occupations (KldB). First, we consider individuals with a university degree and construct a variable that equals 1 if the respondent works as a physician or pharmacist and 0 otherwise. Second, we construct an analogous variable for intermediate educated individuals, i.e. individuals with a secondary school certificate (German: Realschulabschluss) which usually requires 10 years of schooling. For these individuals the variable indicates whether an individual works in a health-related sector, for instance as a nurse, or in another occupation.[53] In the data used, 6 % of the university graduates are working as physician or pharmacist, and 7 % of the intermediate educated individuals pursue a profession within the health sector (see Table 3.1).

Illustrated for the group of academic individuals, the empirical approach is as follows. We compare prevalences in smoking initiation and cessation between physicians or pharmacists and academics working in other occupations. Focusing on the completed

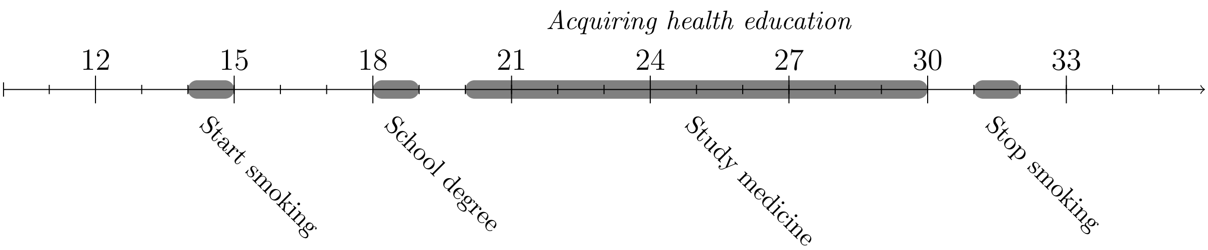
[51] Maralani (2013) finds a similar pattern based on US data and demonstrates formal relationship between the different measures of smoking.

[52] Between 2006 and 2008 all German states passed a smoking ban at schools and school grounds which might have also affected the students' smoking decisions. However, as the most recent wave of our data is from 2009, this law might have rather affected younger cohorts compared to those included in this study.

[53] We consider occupations with the KldB-1992 code 85 including the following professions: nurse, physiotherapist, masseur, midwife, nutrition consultant, alternative practitioner, physician's assistant, medical technical assistant, pharmaceutical technician, speech therapist.

university degree enables us to compare individuals that already finished their studies and thus their health education. Smoking cessation is measured as an indicator variable that equals 1 if an individual reported to no longer smoke and 0 for current smokers.[54] The benefit of comparing equally educated individuals is that we are able to eliminate the selection into education which might also drive the differences in smoking cessation and initiation between the two groups. Differences in the quit rates between physicians or pharmacists and other academics can be interpreted as raw differences due to received health education. Preferences unrelated to the educational decision might also be relevant. We argue that if a specific health knowledge would induce individuals to stop smoking in a causal sense, individuals working in academic health professions should be more likely to stop smoking after the age of 20. The age of 20 is crucial for our analyses as individuals usually take up their (medical) studies at that age and thus begin to receive occupation-specific (health) education. To illustrate the idea, take a 35 year-old man who started to smoke at the age of 14 (see Figure 3.8). As it is common in Germany, he finished his post-compulsory schooling at the age of 19. From the moment the man took up his medicine studies at the age of 20, he receives health education.[55] His decision to take up smoking is made before the health education can be effective and might thus be attributed to selection factors. In contrast, the decision to stop smoking is made after completing his medical studies, at the age of 31. Hence, his decision to stop smoking could be attributed to a causal effect of health education. Unlike the example, the German Microcensus data lack information on the exact age when an individual stops smoking. We thus draw on the information whether an individual used to smoke but reported to no longer smoke in the given survey year.

Figure 3.8: Fictional life course (academic): Acquiring health education



We perform analogous analyses for intermediate educated individuals, i.e. individuals who received a secondary school leaving certificate, distinguishing between health-related occupations, e.g. nurses or physiotherapists. For intermediate-educated individuals, we

[54]The data lack information on age at smoking cessation so that we are unable to conduct analyses based on the individuals' age-specific smoking status as for instance de Walque (2010) or Fuchs (1982) do.

[55]The data lack information on the age when individuals initiate their studies. For that reason, we rely on this early age as it is unlikely that individuals begin their studies at an age before 20 within our sample period.

consider the age 16 to be crucial, as individuals usually take up their vocational training at that age.

Results

Table 3.5 shows the gender- and occupation-specific incidences to start and stop smoking.[56] We first focus on university graduates, i.e. we compare physicians/pharmacists with other academics as presented in Panel a).

Column 1 and 3 present the proportions of smoking initiation across occupations for men and women.[57] The results suggest that individuals who become physicians or pharmacists later in life were already less likely start smoking before the age of 20, i.e. before the individuals receive health education. Among the male physicians and pharmacists (Column 1), 27 % start to smoke at the age of 19 or younger while 31 % of academic men in other occupations initiated smoking before the age of 20. In other words, men becoming physicians or pharmacists later in life have a 4.2 percentage point smaller probability to take up smoking before they receive health education. While the difference in these shares is very similar for academic women (4.1 percentage points), the proportions to take up smoking are again smaller compared to men (Column 3). The differences for both men and women, turn out to be significant at the 95 % level.[58]

Among those who initiated smoking before their studies, the received health education might induce individuals to stop smoking. Of those men who started to smoke before the age of 20, 60 % of the physicians or pharmacists report to have stopped smoking (Column 2). For men working in other occupations this proportion is only slightly lower and amounts to 59 %. Taking the difference, men who finished their medical diploma have a 1.3 percentage point higher probability to quit smoking compared to men working in other academic occupations. For women, the results are again very similar (Column 4). The differences in the occupation-specific probabilities to stop smoking for both, men and women, turn out to be indistinguishable from zero at the 95 % level. However, it has to be considered that the standard errors become relatively large which might be due to the reduced sample size as these analyses are conditional on university-educated individuals that ever smoked. The results nevertheless suggest that health education via medical studies, if at all, has a negligible effect on smoking cessation among university educated individuals.

[56] Estimations obtained from OLS models controlling for German nationality, state fixed effects and a 4th order polynomial in age lead to very similar results (see 3.A.3) We prefer presenting the prevalences and its calculated differences.

[57] We collapse all smoking initiation ages up to age 20. Due to the smaller sample sizes leading to imprecise estimates, we refrain from estimating age- and cohort-specific hazard rates as performed in the previous analyses on educational differences in smoking initiation. However, the results when estimating age-specific hazard rates separately for occupations lead to similar conclusions (not shown).

[58] It has to be kept in mind that there are fewer individuals in the health-related occupations leading to larger standard errors for these groups.

Table 3.5: Proportions in start and stop smoking by health occupations

	Men		Women	
	Start smoking	Stop smoking	Start smoking	Stop smoking
a) Conditional on academic education				
Physicians/pharmacists	0.2708	0.5995	0.1894	0.5893
	(0.0067)	(0.0141)	(0.0068)	(0.0195)
Other academics	0.3125	0.5866	0.2305	0.5740
	(0.0017)	(0.0032)	(0.0019)	(0.0047)
Difference	-0.0417**	0.0129	-0.0411**	0.0154
N	78,597	35,135	50,714	16,913
b) Conditional on secondary education				
Health occupations	0.1560	0.3588	0.1032	0.4022
	(0.0089)	(0.0297)	(0.0023)	(0.0116)
Other occupations	0.1366	0.3734	0.0992	0.3593
	(0.0010)	(0.0039)	(0.0008)	(0.0041)
Difference	0.0194	-0.0146	0.0040	0.0429**
N	108,959	62,313	152,454	65,858

Note: Proportions in start smoking are defined before the age of 20/16, i.e. before health education is received; Proportions in stop smoking are conditional on taking up smoking before the age of 20/16; Standard errors in parentheses; ** p<0.05; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

In contrast to the findings for academics, the results for intermediate educated individuals are not that clear-cut (Table 3.5, Panel b). The shares of taking up smoking before the individuals usually begin their vocational training at the age of 16 are even slightly higher for individuals working in health-related occupations (1.9 percentage points for men and 0.4 percentage points for women). However, these differences are indistinguishable from zero. Regarding the decision to stop smoking, the results are twofold. While men working in health-related occupations are even less likely to quit smoking (1.5 percentage points), women who received occupation-specific health education have a 4.3 percentage point higher probability to stop smoking. This difference is significant for women only, although it has to be mentioned that the standard error for men is relatively large as few intermediate-educated men work in the health-sector. For women, the difference might thus be attributed to the received health education. Alternatively, factors that are simultaneously related to the choice of occupation, such as an health conscientiousness might also account for these differences. The results for intermediate-educated individuals are ambiguous but suggest for women, that the acquired post-schooling health education might induce individuals to stop smoking in a causal sense. In contrast, unobserved characteristics that determine the choice of occupation on the one hand and are related to smoking behavior on the other seem to largely account for the observed differences for university educated individuals. Except for intermediately educated women, we conclude that a specific health education seems to be of minor importance for smoking decisions.

The analyses on health education raise some issues. First, there might be individuals who abandoned their medicine studies before receiving a degree but completed another (non-medical) study. This might bias our results as those individuals received some health education and might be more likely to quit smoking. However, we argue that this bias is negligible as in Germany, the dropout rate for medicine is below 10 % and thus very low (Heublein et al., 2012). Second, the results rest upon the assumption that individuals quit smoking after the age of 20 or 16, respectively, as we are unable to determine the age when individuals stop smoking. This might bias our results if individuals in health-related occupations were already more likely to stop smoking before they take up their medical studies or vocational training. Analyses based on the German Socio-economic Panel[59] suggest that only 8 % among women and 5 % among men decided to stop smoking before the age of 20 (see Figure 3.A.2 in the Appendix). If individuals with health education were nevertheless more likely to stop smoking before this health education was received, e.g. due to an unobserved health consciousness, the observed prevalence in stop smoking after the age of 20 would be overestimated for this group. In turn, the calculated difference can be interpreted as a lower-bound estimate. Finally, a related concern is that physicians are a selective group of individuals as the admission to medicine studies in Germany is highly competitive and generally favors individuals with the highest school grades. Arguing in a similar way, this unlikely biases our results regarding smoking cessation, as the quit rates for physicians are very similar to those of other academics. Regarding the analyses for smoking initiation, this fact is in line with our findings. It supports the interpretation that unobserved characteristics likely determine the choice of occupation, i.e. post-schooling (health) education, as well as the decision to start or stop smoking.

3.8 Conclusions

The aim of this study was to explore how much of the relationship between education and smoking is causal. Our results suggest that differences in smoking initiation rates between low and high educated individuals are already apparent before compulsory education is completed. About 85 % (93 %) of the educational differences in smoking uptake among men (women) are determined before the age of 16. If an individual ever smokes is thus predominantly determined at an age before education is likely to be effective. Differences in smoking initiation rates even diminish over the life course, i.e. the longer individuals have been in school. Completed education might rather be understood as a proxy for

[59]See Wagner et al. (2007) for information on the Socio-economic Panel study. Data for years 1984-2013, version 30, SOEP, 2015, doi:10.5684/soep.v30.

resources gained earlier in life (Maralani, 2013, 2014; Andersson and Maralani, 2015; Jensen and Lleras-Muney, 2012; Pampel et al., 2015).

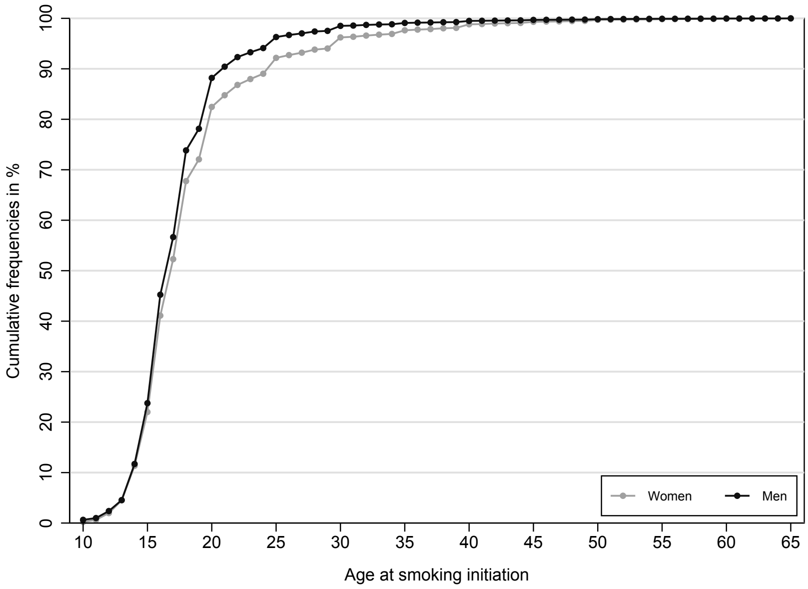
Our results further reveal that mechanisms operating during secondary school, such as peer smoking, account for a large part of the educational differences. This is in line with the limited evidence on the mechanisms during school that highlight the importance of social skills and peer effects (Jensen and Lleras-Muney, 2012; Maralani, 2013; Hsieh and van Kippersluis, 2015). Health education (proxied by health-related occupations) seems to be of minor importance for smoking decisions. However, we find some indication that post-schooling health education induces women with a secondary school leaving certificate to stop smoking.

Previous studies obtained mixed results regarding the causal effect of education on smoking. Thinking about peer effects and other (resulting) mechanisms operating during school might also improve our understanding of this inconclusive evidence. For instance, Jürges et al. (2011) find a strong causal effect of education on smoking exploiting grammar school openings in Germany. Based on the same data source, Kemptner et al. (2011) find little evidence on a causal effect by exploiting changes in the compulsory school attendance. Both results are in line with the explanation that selection and resulting peer effects account for this relationship. While grammar school openings changed the peer group composition in all secondary school tracks, changes in the compulsory schooling only affect the duration of schooling in the lowest track. It does not relax the selection into secondary school tracks.

In general, our results query the external validity of those studies finding a strong protective effect of (college) education on smoking by exploiting some exogenous variation in schooling (e.g. Jürges et al., 2011; Grimard and Parent, 2007; de Walque, 2007). We show that at best, education has very small impacts on smoking – even in a descriptive manner. It is likely that these strong effects are relevant for the complier-specific subpopulation but not for the broader population. This is in line with the recent strand of literature that casts doubt on the causal interpretation of educational differences in smoking or health in general (Farrell and Fuchs, 1982; Maralani, 2013; Clark and Royer, 2013). Further research within the family and school context is required to determine how educational policy could be successful in preventing individuals from smoking. There is a need to rethink the existing theoretical framework to explain educational disparities in smoking.

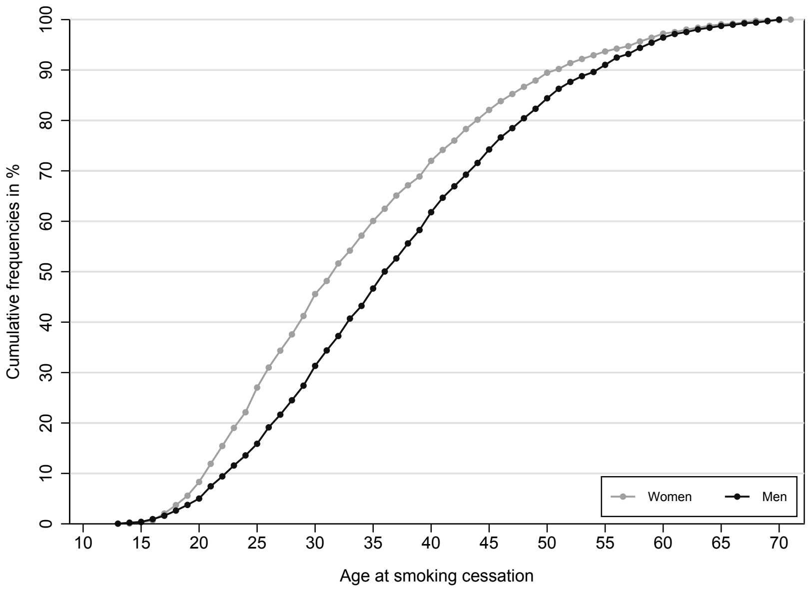
Appendix

Figure 3.A.1: Cumulative distribution of age at smoking initiation (Microcensus data)



Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.2: Cumulative distribution of age at smoking cessation (SOEP data)



Note: Mean age at cessation is 34 for women (N=1,613) and 38 for men (N=2,550). *Source:* SOEP 2002.

Table 3.A.1: OLS estimates: High education (≥ Fach-/Abitur) and ever smoking

	Men		Women	
Cohort	Ever smoking	N	Ever smoking	N
1930–1954	-0.0866***	224,805	0.0098	239,439
	(0.0070)		(0.0119)	
1955–1967	-0.1867***	149,394	-0.1381***	149,468
	(0.0057)		(0.0060)	
1968–1989	-0.2266***	135,440	-0.1667***	137,775
	(0.0070)		(0.0066)	

Note: Controls included: 4th order polynomials of age, i.region*cohort, German citizenship, year; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table 3.A.2: OLS estimates: High education (≥ Fach-/Abitur) and stop smoking

	Men		Women	
Cohort	Stop smoking	N	Stop smoking	N
1930–1954	0.0902***	134,539	0.1294***	76,074
	(0.0073)		(0.0082)	
1955–1967	0.1227***	88,928	0.1461***	74,032
	(0.0047)		(0.0034)	
1968–1989	0.1266****	67,463	0.1310***	56,961
	(0.0045)		(0.0091)	

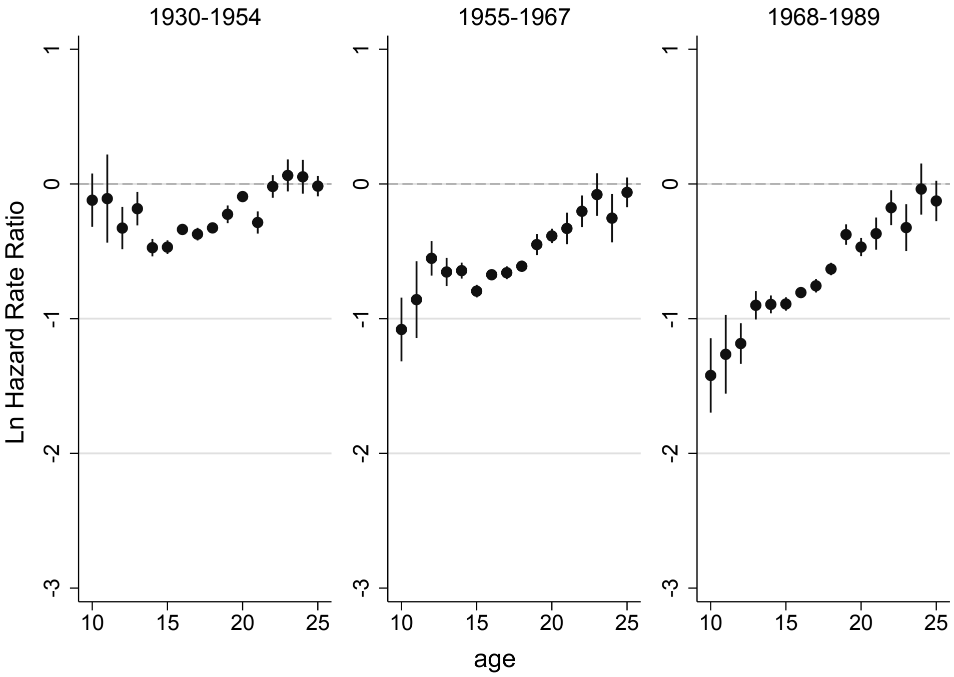
Note: Controls included: 4th order polynomials of age, i.region*cohort, German citizenship, year; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Table 3.A.3: OLS estimates: Start/stop smoking and health occupation

	Men		Women	
	Start < 20/16	Stop smoking	Start < 20/16	Stop smoking
a) Conditional on academic education				
Physicians/pharmacists	-0.0560***	0.0193**	-0.0535***	0.0158
	(0.0084)	(0.0096)	(0.0071)	(0.0203)
N	78,597	35,135	50,714	16,913
b) Conditional on intermediate education				
Health occupations	0.0079	0.0175	-0.0207***	0.0385***
	(0.0096)	(0.0133)	(0.0028)	(0.0053)
N	108,959	62,313	152,454	65,858

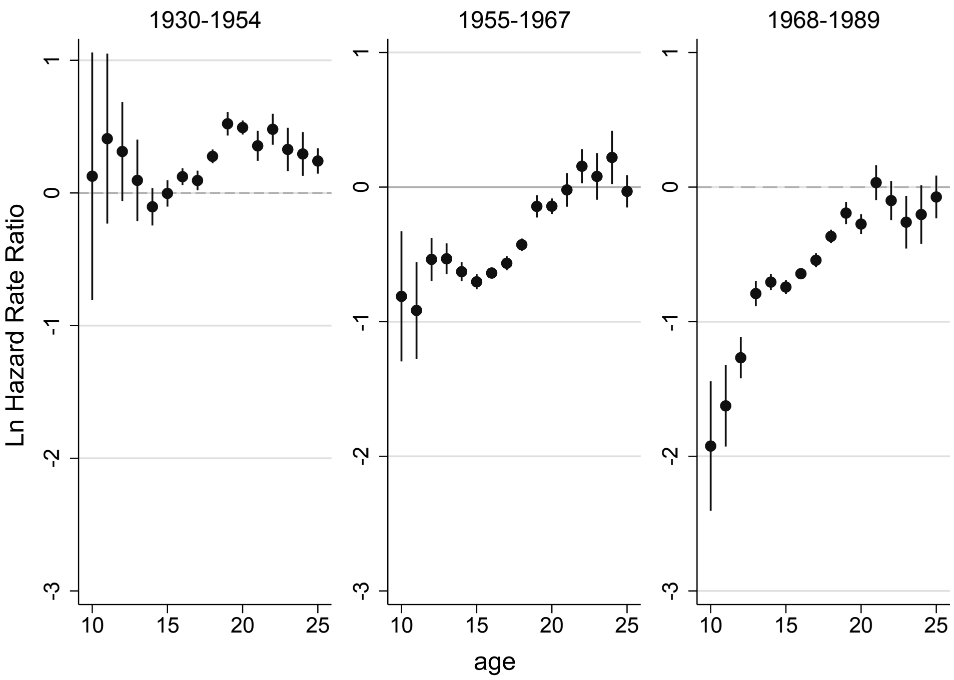
Note: Results for stop smoking are conditional on taking up smoking before the age of 20/16; Control variables included: German nationality, fixed effects for states of residence, fourth order polynomial in age; Robust standard errors clustered at region*cohort level in parantheses; * p<0.10, ** p<0.05, *** p<0.01; *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.3: Log hazard rate ratios: Smoking initiation by completed education, Men



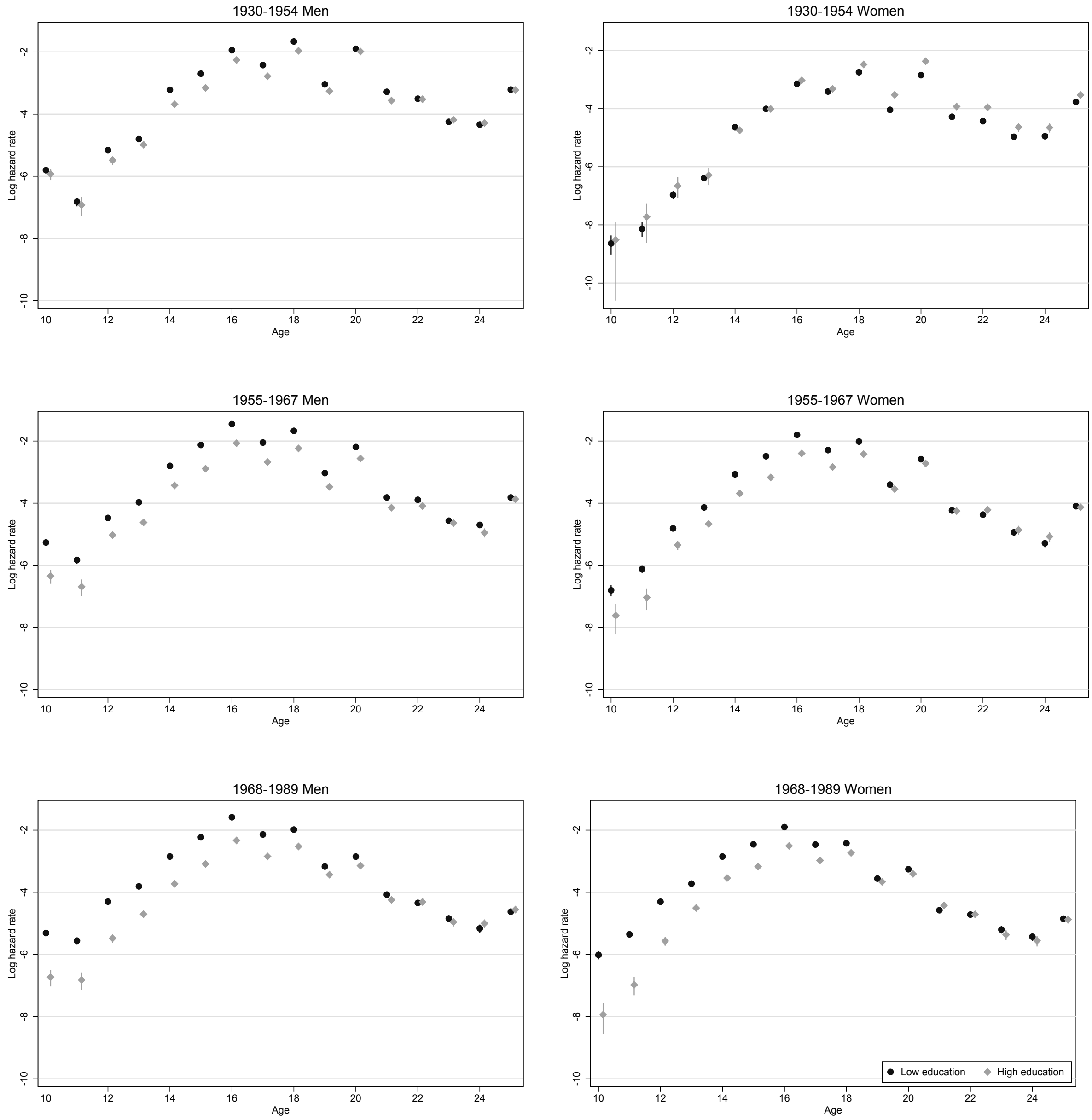
Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.4: Log hazard rate ratios: Smoking initiation by completed education, Women



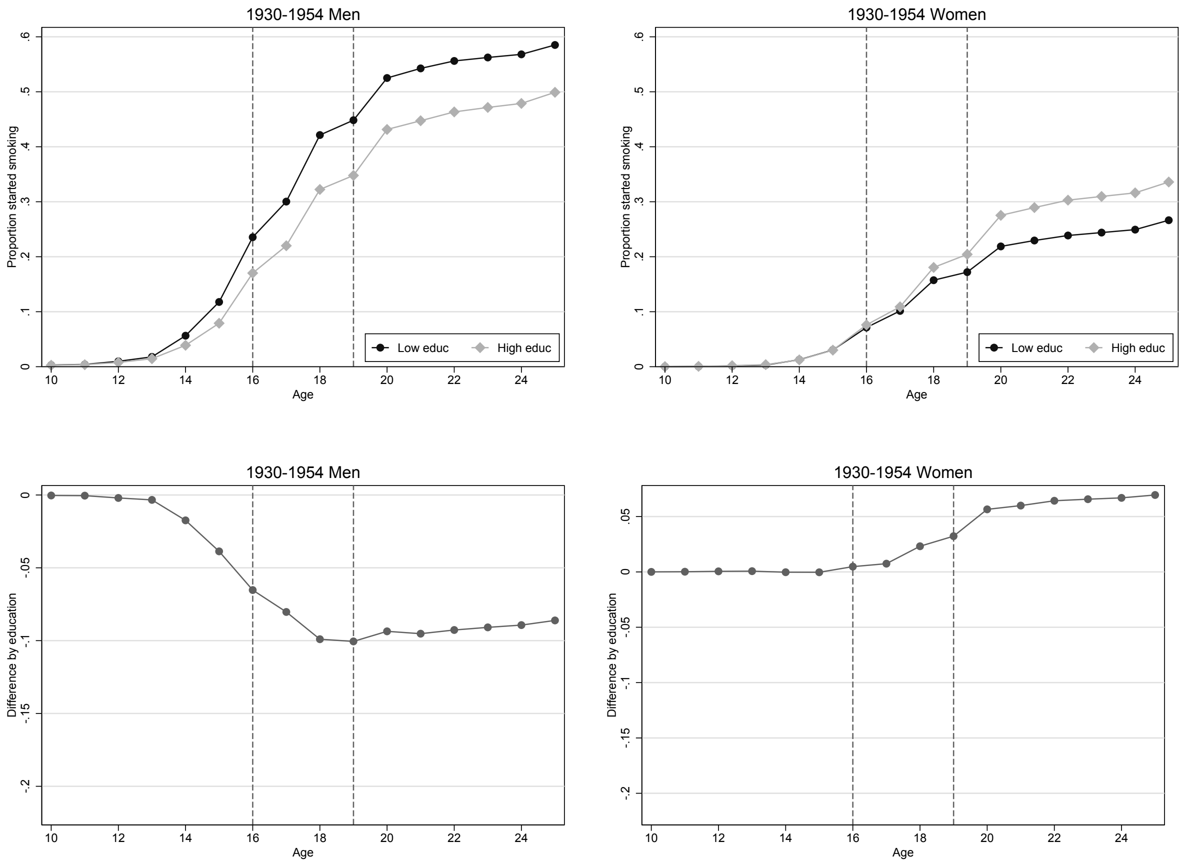
Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.5: Log hazard rates: smoking initiation by completed formal education, East Germany



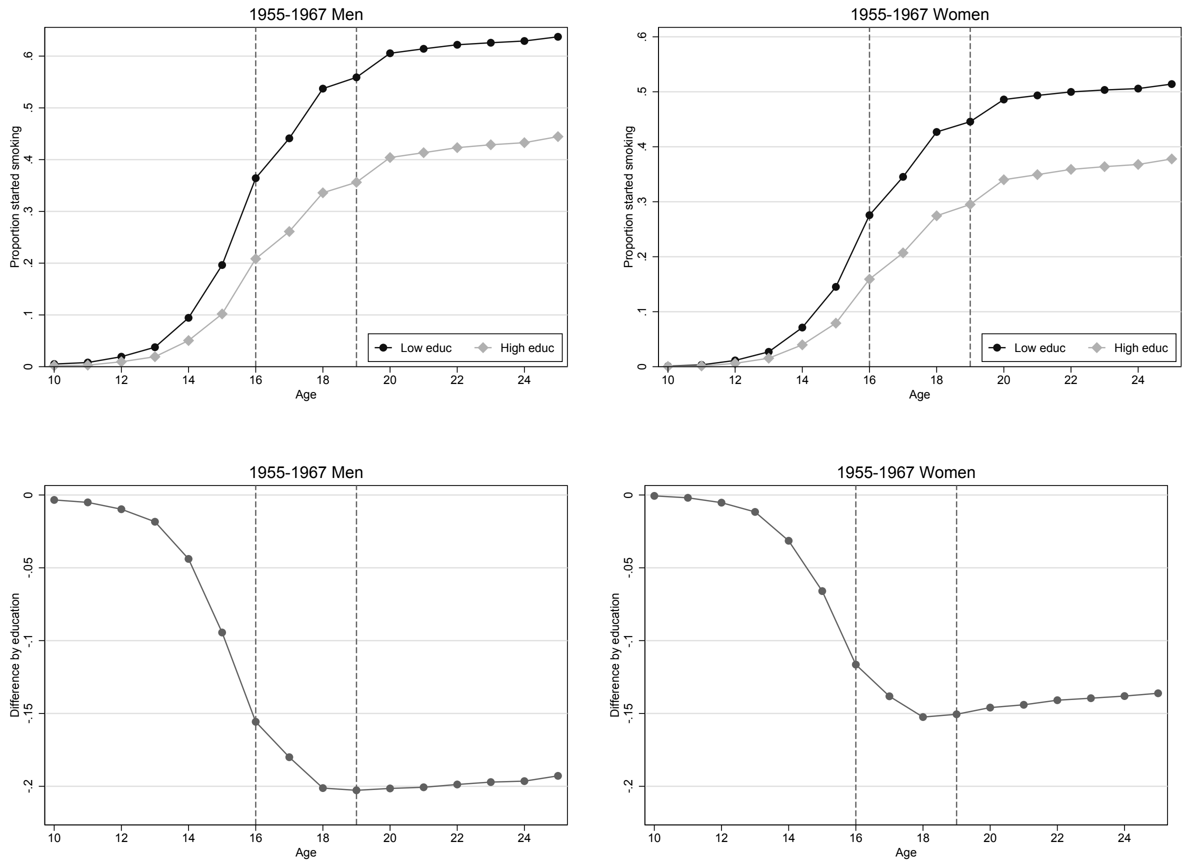
Source: German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.6: Educational differences in smoking initiation until a given age, cohort 1930–1954



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

Figure 3.A.7: Educational differences in smoking initiation until a given age, cohort 1955–1967



Note: Figures in the upper panel show the education-specific distributions while the figures in the lower panel display its calculated differences (higher educated - lower educated); *Source:* German Microcensus 1989, 1999, 2003, 2005, 2009.

CHAPTER 4

Do Occupational Demands Mediate the Educational Gradient in Health (Behavior)?

Joint work with Annemarie Künn-Nelen

4.1 Introduction

There is no doubt that education is related to health. Many studies in various research disciplines have identified a strong relationship between education and a number of health outcomes and the key finding is that the better-educated individuals report better health and face a lower mortality risk (see for instance Grossman, 2006; Cutler and Lleras-Muney, 2006; Mirowsky and Ross, 2003 for literature reviews). In addition, some studies studies investigate the causality of this relationship (e.g. Lundborg, 2013; Grossman, 2006) and find evidence of a causal effect at least for men (Kemptner et al., 2011) and for older ages (Brunello et al., 2016).

It is nevertheless not clear what the underlying mechanisms are. Several studies suggest health behaviors or risky behaviors as mechanisms for the observed educational gradient in health. However, whereas health behavior is indeed related to both education and health (e.g. Cawley, 2011) the relationship is probably driven by economic and time preferences (Fuchs, 1982). This suggests that not health behavior per se is the mechanism between education and health, but the underlying preferences. Moreover, and maybe as a consequence, health behavior does not fully explain the relationship between education and health (Cutler and Lleras-Muney, 2006). Therefore, recent evidence suggests that other mechanisms such as cognitive skills, personality or labor market conditions could mediate the relation between education and health (e.g. Cutler and Lleras-Muney, 2006; Conti and Hansman, 2012; Lochner, 2011; Mirowsky and Ross, 2003; Schneider and Beblo, 2010).

In this chapter, we focus on occupational demands as a specific mechanism in the education-health (behavior) connection. Focusing on occupational demands as possible mediators might be promising as they are likely related to both education and health. On the one hand, education affects the occupational choice due to formal requirements and thereby occupational tasks and demands (e.g. Monden, 2005). On the other hand, the choice of occupation might be interpreted as an investment in the individual's health (Cropper, 1977). Occupational demands are related to workers' health (e.g. Borg and Kristensen, 2000; Fletcher et al., 2011) whether directly through physical demands or indirectly through psychosocial demands. In this study, we distinguish two types of physical demands, i.e. ergonomic demands (e.g. working in a standing position) and environmental demands (e.g. working with dangerous substances). Moreover we differentiate three different psychosocial demands: psychological demands (e.g. working under pressure), social demands (e.g. not being supported by colleagues) and time demands (e.g. working in shifts).

Our study is related to two streams of literature. First, some studies analyze the mediating role of specific occupational demands in the relation between socioeconomic status and health. Most of these studies indeed find a mediating role of occupational demands (e.g. Sekine et al., 2009; Huisman et al., 2008; Kaikkonen et al., 2009). However, as these studies focus on small subgroups of the population it is difficult to generalize their results. Second, there are studies that deal with the relation between education and health, testing the mediating role of occupations by using occupational dummies (e.g. Cutler and Lleras-Muney, 2006; Case and Deaton, 2003). Thereby, the question which specific underlying characteristics are responsible for this mediating effect remains unanswered.

Four studies are similar to the approach of the current study in the sense that they analyze whether specific occupational demands can mediate the relation between education and health (Warren et al., 2004; Brand et al., 2007; Monden, 2005; Qiu et al., 2012). However, we distinguish ourselves from these studies in three ways. First, we additionally consider health behavior and thus aim to also explain the relation between education and health behavior by occupational demands. Focusing on educational differences in health behavior is of particular importance as about half of the deaths are attributable to behavioral factors, such as smoking or excessive weight (OECD and EU, 2014; Mokdad et al., 2004; Cutler and Lleras-Muney, 2010). Second, we use a large representative sample of the entire working population. We also include individuals aged 25–65 who (currently) not work. Omitting this group of individuals (as it is usually done) likely biases the results as these individuals might have quit the labor force due to occupational demands with adverse health effects or because occupational demands have delayed health effects. Third, in contrast to these four studies which rely on self-assessed occupational demands, we estimate occupational demands at a very precise occupational level allowing them to differ across gender and age groups. These occupational demands are likely more objective as we are able to get rid of individual characteristics such as personality traits which likely affect both reported health and perceived occupational demands.

We merge two representative data sources, the German Microcensus 2009 data and the German Employment Survey 2005/2006. The German Microcensus data are unique in the sense that they cover a large representative sample of the German population (1 %). The data include information on various topics such as demographics as well as detailed information on the respondent's education, occupation and some aspects of the individual's health. Based on the occupation individuals work(ed) in, their age and gender, we match this dataset with information derived from the German Employment Survey. We draw on this data to get more insights into the job tasks and demands of the German population. Thereby, the German Employment Survey make it possible to construct the different physical and psychosocial occupational demand indices (cf. Kroll, 2011). Even though the combination of these two datasets does not allow for causal interpretations, it does provide us with a unique dataset that enables a mediation analyses on the role of occupational demands in explaining educational differences and health (behavior).

The results indicate that occupational demands mediate educational differences in subjective health status for lower educational levels only. Regarding the health behavior considered, this partial mediation is more comprehensive. Education coefficients on BMI and smoking status significantly reduce up to 21 % and 27 % when the occupational demands are included. Especially social demands seem to be crucial for the relationship between education and health behavior.

The outline of the chapter is as follows: In Section 4.2, we situate our study in the context of the existing literature. The data used is described in Section 4.3 while Section 4.4 describes the empirical approach with regard to the mediation analysis. In Section 4.5, we report and discuss our findings including some robustness analyses. The final Section concludes.

4.2 Background

We first discuss the literature dealing with the relation between occupational demands on the one hand, and education and health on the other hand. These relations are necessary for occupational demands to be able to mediate the relation between education and health outcomes. Occupational requirements are often related to both the level and field of education. For example, one needs a medical diploma (a masters degree at least) to become a general practitioner. The relationship between education and occupational demands has been shown by a few studies (e.g. Monden, 2005; Borg and Kristensen, 2000). In general, they find that a higher educational level is protective against adverse physical working conditions while it is conducive to psychosocial demands. This suggests heterogeneity in the relationship between education and working conditions across the type of occupational demand.

There is also some evidence on the relationship between occupations and health. For instance, several studies find a faster deterioration in health for manual workers (e.g. Morefield et al., 2011; Choo and Denny, 2006). However, while differences in health between occupational statuses have been well explored, only a few studies put emphasis on more disaggregated occupational groups that operationalize job characteristics rather than social prestige. One reason might be that such analyses require rich data containing enough observations per occupational group to achieve reliable results. A US longitudinal study by Johnson et al. (1999) considers 69 different occupational groups finding that the higher the intensity of work and the less qualified the work, the higher the mortality risk of an occupational group. Even less studies focus on the relationship of specific occupational demands on health. Within a longitudinal setting, Fletcher et al. (2011) find that individuals working in jobs with high physical demands experience significant declines in their health. Regarding psychosocial job demands, most studies rely on sociological stress theories according to Karasek (1979) or Siegrist (1996). According to these theories occupations can be classified by their extent of job demand (effort) in relation to the extent of job control (reward). In general, the literature relating to these theories has found that an imbalance of demand and control or effort and reward is associated with worse health (see e.g. Bakker and Demerouti 2007 for an overview). However, these approaches rather focus on one psychological and largely subjective aspect of occupations than mapping the work environment of occupations more broadly.

Previous literature suggests that occupational demands may indeed mediate the relationship between education and health. Several studies analyze the mediating role of specific occupational demands in the relation between a broad measure of socioeconomic status and health. Kaikkonen et al. (2009) examine the role of different physical and psychosocial working conditions in mediating occupational inequalities in self-rated health among municipal employees of the City of Helsinki. They find that physical working conditions account for about half the inequalities among men, while job control mediates about 40 percent of the inequalities among women. A different study investigates the link between occupational demands and socioeconomic inequalities in the incidence of myocardial infarction (Huisman et al., 2008). This study finds that job control and adverse physical working conditions provide a partial explanation of the relationship. Additionally, Sekine et al. (2009) conclude that adjusting for working conditions moderately attenuates the socioeconomic differences in poor mental and physical functioning among civil servants from Britain, Japan and Finland. Using French data, Niedhammer et al. (2008) find that ergonomic, physical and chemical exposures, as well as self-perceived decision latitude, reduce the occupational class differences in health by 24–58 %. However, they find no evidence that any other self-perceived psychological demand contributed to the link. In contrast, Lahelma (2004) finds that occupational class mediates the relationship between education and health only to a very small extent. This different finding might again be explained by the fact that occupational class rather measures different aspects of occupations, such as prestige, compared to specific working conditions. Except for the last study, all studies find that specific occupational demands mediate the relation between socioeconomic status and health to some extent. However, as most of these studies focus on a small subgroup of the population only, one should be careful in generalizing these findings to the entire working population.

There are also a few studies dealing with the specific relation between education and health. But these studies often test the mediating role of occupations by occupational dummies, instead of specific occupational characteristics, such as occupational demands. Cutler and Lleras-Muney (2006) control for occupation and industry dummies in longitudinal data and did not find that occupations can be the main mechanism by which a higher education relates to better health. Case and Deaton (2003) find that effects of education are reduced though not eliminated by controlling for 16 occupational groups. They conclude that lower-paid work in manual occupational groups impairs self-rated health much stronger and health also worsens even faster over time. However, from these studies, it remains unclear what occupational characteristics are responsible for the mediating effect they identified.

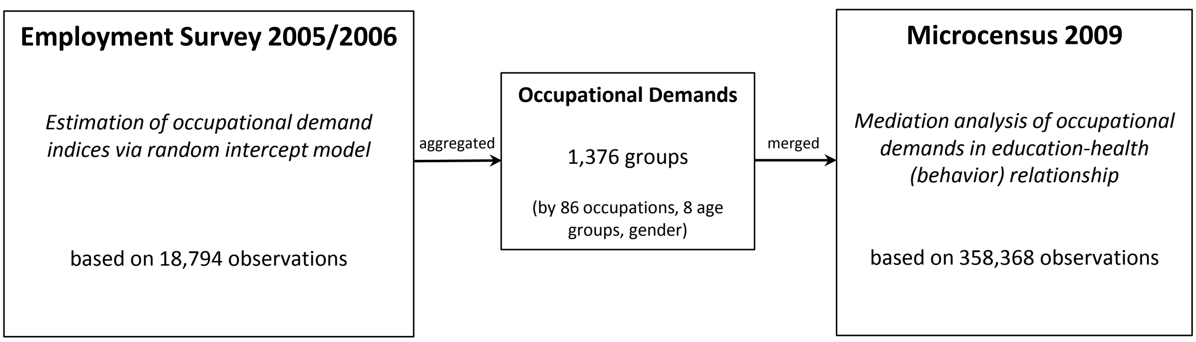
There are four studies explicitly looked into the mediating role of occupational demands in the relationship between education and health. Apart from other measures of socioeconomic status, Warren et al. (2004) also analyze the role of physical and psychosocial job characteristics in mediating the relationship between the probability of a completed college education and health of high school graduates. They find that the occupational demands used account for some or all of the associations between education and health, depending on the health outcome considered. In a successive study, Brand et al. (2007) obtain very similar results restricting their analyses to sibling pairs. However, as they only consider high school graduates, a generalization of their findings is scarcely possible. For the Dutch population, Monden (2005) investigates the mediating role of different physical and psychosocial working conditions in the relationship between education and health. Taking retrospective information, he finds that lifetime exposure to adverse working conditions explains about one third of educational differences in subjective health for men but only a small percentage for women. Qiu et al. (2012) examine the relationship between psychosocial working conditions and educational disparities in health and find heterogeneity in the associations across different types of demands, distinguishing psychosocial resources and demands. In this study, we add to these four studies by focusing not only on subjective health status but also on health behavior. Moreover, we use a large representative sample of the working population, including individuals who no longer work. The reason for doing so is that omitting the non working would bias the results as they might have quit working due to occupational demands with adverse health effects.

Theoretically, several pathways for how occupational demands might operate as mediators in the education-health (behavior) relationship are conceivable. With respect to physical demands it is likely that the lower educated choose or depend on jobs with poor working conditions, which could affect workers health directly due to attrition or accidents. Moreover, it is also possible that physical demands promote unhealthy behavior. People working in physically demanding occupations are likely to have an increased need for periods of rest during the working day. Smoking could be used as an excuse for taking such short breaks (Albertsen et al., 2004). In addition, smoking breaks could be culturally-rooted, especially in manual occupations. As a result, peer effects are likely to occur, as workers might be more inclined to smoke or also eat if their co-workers do so. Unlike physical demands, psychosocial demands probably occur in both low-skill and in high-skill occupations. While the lower educated could experience psychosocial demands such as a small degree of autonomy at work, the higher educated are more likely to experience a different kind of psychosocial demand, e.g. becoming new tasks often or working overtime (e.g. Qiu et al., 2012). Both could be perceived as stressful, and it is therefore not a priori clear how education might be related to psychosocial demands. However, compared to high-educated workers, it is possible that the low-educated workers have a lower level of (health) knowledge and poorer coping strategies to deal with psychosocial demands, e.g. time pressure or lack of social support. In turn, this could influence their health behaviors, as smoking or excessive eating could serve to compensate for such stressful demands.

4.3 Data and Measures

To explore the mediating role of occupational demands in the education-health (behavior) nexus, we make use of two different data sources. First, the main data source is the German Microcensus 2009. The German Microcensus is an annual administrative cross-section covering approximately 1 % of the German population with interviews imposing a duty of disclosure. As these data lack detailed information on working conditions, we draw on a second data set to construct the occupational demand indices, the German Employment Survey 2005/2006. The German Employment Survey is a labor force cross-section on qualifications and working conditions in Germany covering 20,000 employees (Hartmann, 2006). The survey includes detailed information on physical and psychosocial occupational demands. Both datasets are representative for the German population. As these data sources are two independent surveys it is not possible to identify individuals across these datasets. We match the occupational demands via 86 occupations (KldB 2-digit) but also gender and eight age groups to the German Microcensus 2009. The variation is extended as occupational demands not only differ by occupations but also by gender and age. The procedure is illustrated in Figure 4.1. The data offer enough cases, i.e. there is one separate value for each combination (1,376) of the merge variables. The construction of the occupational demands is described in Section 4.4.1.

Figure 4.1: Merging occupational demands to the Microcensus



The mediation analyses are based on this matched dataset. We restrict the analyses to individuals aged 25–65 in order to concentrate on the working-age population and on those who have likely completed their education. Within the German Microcensus, we only consider individuals living in private households at their main residence in order to avoid some individuals to be counted twice. Our final sample consists of 358,368 individuals who participated in the German Microcensus. We include working and currently not working individuals in our analyses. For currently not employed individuals, for example unemployed or (early) retired individuals, we rely on their last occupation. We thus assume that former working conditions are related to their current health (behavior).[60] We include non-working individuals as there could be a non-random selection of workers who have left the active labor market because occupational demands (in their last job) negatively affected their health and they are no longer able to work.[61] The working population might therefore be a positive selection of people, which is defined as the so-called healthy worker effect (McMichael et al., 1974). For that reason, we construct a continuous variable that indicates how many years have elapsed since the respondent has left his/her last job. To distinguish the non-working from the working population, this variable equals 0 if the respondent is currently employed. As we merge the occupational demands via occupational groups (age group and gender), we exclude individuals without information on their last or current occupation and with missing information on the main variables.[62]

Main Variables

We consider self-assessed health status and health behavior as outcomes. The question on health status (*Have you been ill or injured during the last four weeks?*) is translated into a dichotomous variable taking the value 1 if the respondent has been in bad health during the last four weeks and 0 otherwise. Given this specific question our health measure rather captures short-term illnesses during the last month than long-term or chronic diseases.

[60] For the employed individuals in our sample, we have to assume that occupational demands did not substantially change in the years between the dates of the two surveys (2006 and 2009). Unfortunately, the German Employment Survey provides data on occupational characteristics for 2006 only. Thus, for the no longer employed in 2009, we have to assume that occupational demands (related to their last jobs) have not changed significantly since they have left their jobs. The average time span between 2009 and the withdrawal from the labor market is 7.8 years. Both samples are fairly similar distributed with regard to the main variables considered in this study (see Table 4.A.1 in the Appendix).

[61] Implicitly, we thus make the assumption that past (cumulative) exposure to working conditions matters. Indeed, cumulative exposure to work factors seems to be important as persistent occupational demands affect health in the long term to a greater extent (see Fletcher et al. 2011 or Monden 2005). However, we are not able to test this as the data lack detailed information on work histories and are cross-sectional. Additionally, we are not able to consider job transitions which might occur due to harmful occupational demands.

[62] This concerns 9,649 people who have either never worked or did not report valid information on their occupation.

Our measures of health behavior are current smoking and body mass index (BMI).[63] These outcomes are of particular relevance as smoking and obesity are the dominating behavioral factors that promote various health problems and even death (OECD and EU, 2014; Mokdad et al., 2004). Current smoking is used as an indicator variable, taking the value 1 if the respondent occasionally or regularly smokes and 0 otherwise. BMI is calculated by the common formula (weight in kg/height in m^2), using self-reported height and weight. We consider BMI as a continuous outcome in the analyses.[64] It has to be mentioned that BMI is rather a health stock measure than a health behavior. In the literature, it is commonly used as an indicator for future health or risky health behavior (e.g. Kemptner et al., 2011). As the data lack information on direct health behavior other than smoking, we interpret BMI as a proxy for (past) dietary and activity habits.

As the Microcensus contains separate information on the highest education level and the type of educational degree (general versus vocational), a combined variable considering both educational level and the type of educational degree is constructed according to the CASMIN classification. The CASMIN classification is one of the most frequently applied instruments to measure education in an international comparative sense (Brauns et al., 2003). Education is classified into eight different stages: primary education (if the respondent has not completed any kind of secondary schooling), lower, intermediate and upper secondary education each combined with the information whether the respondent has completed any kind of vocational training (voc) or not (gen). The highest educational group consists of individuals with tertiary education and serves as reference category in our analyses.

Occupation is measured by the classification of professions (1992) of the German Federal Employment Agency's (KldB92) 2-digit codes, distinguishing 88 different occupational groups. As the Employment Survey 2005/2006 lacks observations on two occupational groups we distinguish 86 different groups in the analyses. The German classification is very similar to the International Standard Classification of Occupations (ISCO) but structures occupations more adequately for the German labor market.[65]

We include gender, age, age^2, marital status, region (West vs. East Germany) and the number of hours usually worked by the respondent per week as covariates in all our regressions.[66] As there is no information on the number of working hours in the

[63]There are limitations with the measurement of these health outcomes as they are based on self-reported data and the results could suffer from measurement error. Measured BMI is generally higher than self-reported (Cawley, 2004). Thus, BMI as well as smoking could typically be underestimated for example due to social desirability.

[64]BMI as a measure of overweight is often criticized due to the fact that muscle mass leads to a higher BMI (Burkhauser and Cawley, 2008).

[65]For instance, group 78 consists of office clerks, group 11 of occupations in agriculture, forestry, and farming. See `https://www.destatis.de/DE/Methoden/Klassifikationen/Berufe/KlassifikationKldb92` for details on this classification

[66]The number of hours worked is likely to be endogenous to health status, particularly to short term illnesses and disabilities. Against the background that our study makes no claims of being causal, we

respondent's former occupation and thus for the no longer employed individuals, we impute them by taking the mean working hours per week via gender, age group (when occupation was left) and the respondent's last occupational group.

4.4 Empirical Approach

In this chapter, we investigate the mediating role of five different types of occupational demands in explaining the relation between education and health. The occupational demands are calculated via a random intercept model based on the German Employment Survey. The predicted occupational demands derived from this model are aggregated across occupations, age groups and gender and merged to the German Microcensus data. On the basis of this combined dataset we explore the mediating role of occupational demands in the education-health (behavior) relationship. This approach is explained in more detail in the following subsections.

4.4.1 Construction of Occupational Demand Indices

We build on Kroll (2011) by performing a multilevel analysis with random intercepts to generate occupation-specific demands. This approach to measure occupational demands deviates from the way that is commonly used in occupational epidemiological research, namely calculating so called Job Exposure Matrices (JEM) by taking occupation-specific means. However, relying on JEM is problematic, as it implicitly assumes that the observed characteristics result from the features of the occupation only and that there are no other important differences between the workers. Thereby, group and individual effects on the outcomes cannot be disentangled. However, the applied method in this study is also different from previous cross-sectional research on mediating effects of occupational demands in the education-health relationship, which relies on self-assessed occupational demands (e.g. Warren et al., 2004; Monden, 2005). Including self-assessments of occupational demands rather captures personal characteristics than objective occupational demands to some extent. We therefore pursue another approach and assume that occupational demands themselves are mainly driven by differences in occupations, age, working hours per week and gender.

Calculating occupational demands via a multilevel model yields more robust estimates compared to calculating simple occupation-specific means, especially for small occupations. In contrast to Kroll (2011), we consider each of the five categories separately – ergonomic, environmental, psychical, social and time demands – instead of adding

include this variable as there might be substantial differences in the occupational demands according to the number of working hours.

these categories to a physical and psychosocial index.[67] Thus, occupational demands are included in a more detailed way to avoid aggregation bias, as each of them might be differently correlated with education and health. The procedure of generating the five demand indices contains several steps. First, the 39 single items on occupational demands (see Table 4.A.2) collected in the Employment Survey 2005/2006 are dichotomized (having harmful demand often vs. never, seldom, sometimes), assigned to the five categories (ergonomic, environmental, social, psychological and time demands) and added to an individual sum score for each of the five categories. In case that the conditions are health enhancing, e.g. arranging work autonomously as well as the social demand items, the labels are interchanged (never vs. seldom, sometimes, often). These sum scores are z-standardized to account for the different number of single items. Finally, the five individual indices – ergonomic, environmental, psychological, social and time occupational demands – are taken as outcome variables for the random intercept model to calculate the final indices by adjusting for the interceding variables gender, age and working hours:

$$Demands_{i,j} = \beta_0 + u_{0,j} + \beta_x X_i + \epsilon_{i,j}, \tag{4.1}$$

in which $Demands_{i,j}$ denotes one of the five different types of occupational demands of individual i in occupational level $j = 1, 2, 3$. Multilevel regressions with random intercepts cope with the nested structure of data, here individuals nested within occupations, by dividing the overall error term in one separate random error term per level (indicated by subscript j). We exploit the full hierarchical structure of the occupational classification (KldB) by considering three levels (indicated by subscript j): the 2-digit (86 different groups), 3-digit (369 groups) and 4-digit (2,287). The occupational level specific intercepts ($u_{0,1}, u_{0,2}, u_{0,3}$) represent unobserved heterogeneity across each occupational level. Thus, the overall variance of occupational demands is separated into variation that is attributable to 2-digit occupation-specific characteristics, variance that is attributable to differences between 3-digit occupations, variance that can be ascribed to the 4-digit occupations and finally the residual variance that is attributable to other, individual-specific characteristics (Rabe-Hesketh and Skrondal, 2008). Hence, it is assumed that occupational demands arise from the sum of parameters for overall job demands (β_0), the occupational group specific demand on the level of KldB 2-digit, 3-digit and 4-digit ($u_{0,j}$), a vector of covariates including gender eight age dummies, five dummies for the usual number of working hours per week ($\beta_x X_i$) and the individual error term ($\epsilon_{i,j}$). To assess the proportion to which the total variance in occupational demands can be ascribed to the different occupations in their different levels, variances corresponding to the different levels and consequential the intra-class correlations are reported. Intra-class correlations

[67]There are further differences between Kroll's (2011) approach and ours: Job tenure is not included as covariate because there is no information on it in Microcensus. Additionally, we include the control variables in categories.

represent the within cluster correlation, i.e. the share of variance attributable to the occupational level that ranges from 0 if the grouping conveys no information to 1 if all members of a group are identical.

We exclude individuals with missing information on the occupational demands items. The analysis sample of the Employment Survey amounts to 18,794 working individuals. Table 4.1 presents the results of the random intercept models. The high intra-class correlations suggest that the different occupational demands considered vary substantially between the three different occupational levels (see Table 4.1 in the Appendix). Especially ergonomic and environmental occupational demands vary substantially between occupations, to that extent, that the variation attributable to all occupational levels is even greater (51 % and 57 %)[68] than the residual variation, representing the variation on individual level. Differences in social and time occupational demands (15 % and 30 % respectively) can be ascribed to the occupational levels to a lower extent. The variation in psychological demands is merely attributable to occupational groups (6 %). This could be due to the fact that the nature of psychological demands is rather subjective.[69]

Table 4.1: Random intercept results: Generating occupational demands

	Ergonomic	Environm.	Psychol.	Social	Time
Intra-class correlations					
KldB_2-digit	0.387	0.394	0.022	0.105	0.131
KldB_3-digit	0.098	0.056	0.014	0.030	0.089
KldB_4-digit	0.083	0.064	0.024	0.019	0.076
Chi^2	42.2	256.5	1410.8	155.7	3696.1
Variances					
Var KldB_2-digit	0.447	1.321	0.100	0.129	0.263
Var KldB_3-digit	0.113	0.189	0.062	0.037	0.179
Var KldB_4-digit	0.095	0.216	0.107	0.023	0.154
Var residual	0.498	1.629	4.185	1.037	1.415
LL	-20951.1	-31947.7	-40357.3	-27316.7	-30541.4
N	18,793	18,793	18,794	18,791	18,794

Note: All models include gender, age group dummies (25–30, 31–35, 36–40,. . . , 61–65), dummies for working hours/week (10–20, 21–30,31–40, 41–50, >50) as covariates; * p<0.05, ** p<0.01, *** p<0.001; *Source:* Employment Survey 2005/2006.

Based on these models, we predict the occupational demands in a next step, considering the fixed and random part of the random intercept model. Finally, the five demand indices are aggregated by gender, 8 age groups and 86 occupations (KldB 2-digit) and

[68]This is the result when all intra-class correlations of the three different occupational levels are taken together: e.g. ergonomic demands: 0.394 + 0.056 + 0.064=0.51.

[69]See Table 4.A.3 for a correlation matrix between the different occupational demand indices.

merged to the German Microcensus 2009. We z-standardize the demand indices for the ease of interpretation.

4.4.2 Mediation Analyses

Based on the merged dataset we finally perform the mediation analyses. We follow the procedure suggested by Baron and Kenny (1986) to investigate whether any of the considered occupational demand indices mediates the relationship between education and health or health behavior, respectively. We use ordinary least squares (OLS) models with robust standard errors clustered at the occupation level (KldB 2-digit) in all models.[70] In doing so, we take the aggregated nature of the occupational demands into account and render the estimates' significance levels comparable across the estimated models, also in models where occupational demands are not included.

We first estimate the baseline relationship between the different education ($educ_i$) and the health outcomes bad health, BMI and smoking ($Health_i$):

$$Health_i = \gamma_0 + \gamma_1 Educ_i + \gamma_x X_i + \omega_i, \tag{4.2}$$

where the vector X includes the control variables discussed in Section 4.3. We also include the number of years since the respondent has left his/her last occupation which equals 0 for the currently employed. That way, we are able to distinguish the currently employed and the currently not employed subpopulations.

Second, we estimate one of the conditions that have to be fulfilled when testing mediations: We regress ergonomic, environmental, psychological, social and time occupational demands on education to evaluate whether occupational demands are significantly related to the different educational level dummies. The occupational demand indices ($Demand_i$) generated via the multilevel approach as described in Section 4.4.1 are used as dependent variables in Equation (4.3):

$$Demand_i = \delta_0 + \delta_1 Educ_i + \delta_x X_i + \zeta_i. \tag{4.3}$$

Third, we estimate the relationship between occupational demands and health which is the second condition that should hold for occupational demands to mediate the relationship between education and health:

$$Health_i = \eta_0 + \eta_2 Demands_i + \eta_3 Demands_i \times Years_i + \eta_x X_i + \psi_i, \tag{4.4}$$

[70]For the ease of interpretation, we estimate linear probability models for the indicator outcomes bad health and smoking. In a robustness check we also estimated probit models for these outcomes in Equations 4.2, 4.4 and 4.5. As the results are very similar, we solely report the LPM. We refrain from performing multilevel analyses as intra-class correlations are very small.

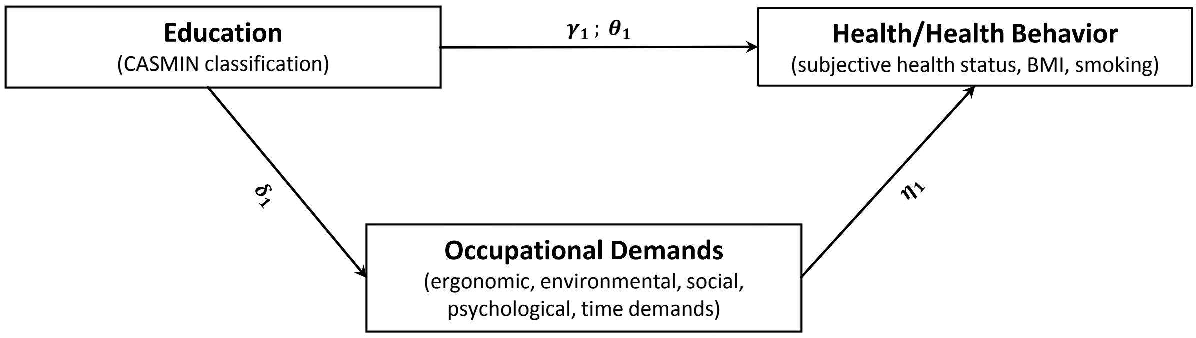
including interaction terms between occupational demands and the number of years since leaving the last job. This enables us to see whether occupational demands are differently related to health across individuals who are currently employed and individuals who are (temporarily) dropped out of employment.

Finally, we estimate the mediation model in which health is regressed on both education and occupational demands:

$$Health_i = \theta_0 + \theta_1 Educ_i + \theta_2 Demands_i + \theta_3 Demands_i \times Years_i + \theta_x X_i + \xi_i. \tag{4.5}$$

In the main analyses, we include all five demands simultaneously. However, we also estimate specifications in which we include each occupational demand separately to assess which occupational demands are most important (see Figures 4.A.5, 4.A.6 and 4.A.7 in the Appendix). Figure 4.2 summarizes the different steps of the mediation analyses taking occupational demands as intervening variables in the education-health (behavior) relationship.

Figure 4.2: Mediation analyses framework



Comparing γ_1 in Equation (4.2) and θ_1 in Equation (4.5), shows whether and to what extent the occupational demands mediates the relationship between education and health (behavior). If the occupational demands considered fully mediated the relationships of interest, the education coefficients would become zero. However, given the rather complex relationship between education and health (behavior) and the findings of previous related studies, we expect occupational demands, if at all, to partly mediate the correlation between education and health. In that case, we expect $\theta_1 < \gamma_1$. In order to assess the extent to which the inclusion of occupational demands changes the education coefficients, we calculate the change in terms of percentages:

$$Change = \frac{\theta_1 - \gamma_1}{\gamma_1} \times 100. \tag{4.6}$$

To assess whether these changes are statistically meaningful, the corresponding standard errors are computed by a bootstrapping procedure drawing 500 samples clustered at the occupation level (KldB 2-digit).[71] The corresponding t-values are reported to test the significance of the difference between the two coefficients. It has to be considered that there might be potential alternative mechanisms, e.g. that individuals sort themselves in certain occupations because of their health limitations. These mechanisms are not addressed in this study and our approach should rather be considered in a descriptive sense. However, a comprehensive explanation of causal mechanisms via mediators is very difficult to address even in experimental mediation analysis (Bullock et al., 2010). Biased estimates likely occur if the direct effect is mediated by more than one variable or if the independent and mediating variables correlate. In order to reduce the resulting bias, randomized experiments are a viable solution. Yet, finding appropriate instruments that externally manipulate the independent (education) and mediating (occupational demands) variable is a difficult task. As these problems are not satisfactorily solved by standard statistical procedures, we do not claim any causal interpretations of our results. In order to assess the potential bias due to the pooling of employed and not employed individuals, we additionally perform a robustness check with separate analyses for these two groups.

4.5 Results

4.5.1 Requirements for the Mediation Analyses

Occupational Demands and Education

Table 4.2 displays the results for Equation (4.3), the relationship between occupational demands and education (see also Figure 4.A.4 in the Appendix). Except for time demands, all occupational demands are significantly related to all educational levels. The results even point to an educational gradient, where lower educated individuals work on average in occupations with more hazardous physical (ergonomic and environmental) and social occupational demands compared to tertiary-educated individuals. The results are most pronounced for social demands: Average social demands within the occupations of primary educated individuals are 1.6 standard deviations larger. These findings are consistent with previous studies (e.g. Monden, 2005). In contrast, psychological demands are negatively related to education with tertiary educated individuals experiencing the

[71] The variance of the difference would be calculated by $Var(\gamma_1 - \theta_1) = Var(\gamma_1) + Var(\theta_1) - 2Cov(\gamma_1, \theta_1)$. However, since the coefficients are derived from two different regressions, the covariance of the two coefficients is not easily available. For that reason, the literature suggests to apply more sophisticated approaches to solve this problem, such as structural equation modeling/seemingly unrelated regressions or a bootstrap procedure (see, e.g., Biesanz et al. 2010 for a discussion).

Table 4.2: OLS estimates: Education and occupational demands

	(1) Ergonomic	(2) Environm.	(3) Psychol.	(4) Social	(5) Time
Educational level					
Primary (1a)	1.485***	1.334***	-1.201***	1.627***	0.156
	(0.196)	(0.130)	(0.241)	(0.298)	(0.209)
Lower secondary gen (1b)	1.338***	1.259***	-1.150***	1.485***	0.164
	(0.197)	(0.138)	(0.208)	(0.252)	(0.198)
Lower secondary voc (1c)	1.035***	1.057***	-0.880***	0.926***	0.083
	(0.197)	(0.159)	(0.151)	(0.178)	(0.177)
Intermediate secondary gen (2a)	0.778**	0.804***	-0.872***	0.966***	0.016
	(0.230)	(0.163)	(0.155)	(0.219)	(0.197)
Intermediate secondary voc (2b)	0.596**	0.659***	-0.622***	0.584***	-0.003
	(0.206)	(0.155)	(0.123)	(0.144)	(0.218)
Upper secondary gen (2c_gen)	0.363*	0.326*	-0.559***	0.473**	0.058
	(0.171)	(0.125)	(0.119)	(0.149)	(0.189)
Upper secondary voc (2c_voc)	0.272	0.324**	-0.377***	0.297**	-0.082
	(0.177)	(0.119)	(0.0993)	(0.105)	(0.225)
Tertiary	*reference*	*reference*	*reference*	*reference*	*reference*
F (all education levels)	12.82***	17.06***	7.18***	8.01***	1.37
Adj. R	0.204	0.342	0.348	0.258	0.223
# of Clusters	86	86	86	86	86
N	358,368	358,368	358,368	358,368	358,368

Note: All models refer to Equation 4.3 and include gender, age, $(age/10)^2$, marital status, West Germany and working hours/week as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

highest psychological demands. The results also indicate an educational gradient in occupational demands, i.e. the associations become successively stronger the lower the educational level. Psychological demands of primary-educated individuals are on average 1.2 standard deviations smaller compared to those of their tertiary-educated counterparts. Except for time demands the F-statistics on joint significance also suggest that education as a whole is significantly related to the occupational demands.

Occupational Demands and Health (Behavior)

For occupational demands to mediate the relation between education and health, occupational demands should also be related to each of our health outcomes. In a next step, we thus estimate the relationship based on Equation (4.4). The raw correlations are illustrated in Figures 4.A.1, 4.A.2 and 4.A.3 in the Appendix. We find that all occupational demands are significantly related to health status, BMI and most strongly to smoking.

The only exception are again time demands, which are not significantly related to any of the health measures.

Table 4.3 shows the relation between occupational demands and the three different health outcomes conditional on a range of covariates (see Section 4.3). We include interaction terms between the occupational demands and the number of years since leaving the last job (see Section 4.4). For that reason, the coefficients for the occupational demands show the associations for employed individuals, i.e. if the number of years since the last job are zero. For the currently not employed individuals the relationship is indicated by the sum of occupational demand coefficients and the corresponding interaction.

Regarding bad health (Column 1), we find weak correlations for the currently employed, turning out to be significant for social demands only. The positive correlation implies that the higher the social demands within an occupation, the larger the employed individual's probability to be in bad health. Additionally, we find a positive relation between the probability to be in bad health and the years since the last occupation. The probability to report bad health increases with each additional year out of the job by 0.5 percentage points. The interaction terms between the years since last occupation and the occupational demands are significant for all demands. For the no longer employed, positive interaction terms reveal that the probability to be in bad health is even larger for individuals that worked in an occupation with high environmental or time demands. In contrast, negative interaction terms suggest that the probability to be in bad health decreases for the no longer employed if they worked in occupations with relatively high social, psychological and ergonomic demands. These strong correlations of the interactions could be a hint that individuals have quit their jobs for health reasons due to hazardous occupational demands. Moreover, occupational sorting patterns might drive the results. We further explore this in the robustness analyses in Section 4.5.3. Overall, the F-test on joint significance indicates that the occupational demands are altogether significantly related to the probability to be in bad health at the 1 % level.

The results for BMI (Column 2) also suggest social demands to play a major role. We observe a significant and positive relation between BMI and social demands for the currently employed, indicating that a standard deviation increase of social demands is associated with a 0.32 kg/m^2 higher BMI. The BMI slightly increases with each additional year out of the job. For currently not employed individuals, the occupational demands are only weakly related to an individual's BMI and do not differ significantly from 0, except for social demands. However, the joint significant test suggests that overall, occupational demands are significantly related to BMI.

Compared to the other health measures, the correlations between occupational demands and smoking are most pronounced. While a standard deviation increase of social demands within an occupation is related to a 2.5 percentage point higher probability to smoke for employed individuals, higher psychological demands seem to decrease an indi-

Table 4.3: OLS estimates: Occupational demands and health (behavior)

	Bad health	BMI	Smoker
Occupational demands			
Ergonomic	0.0042	0.0736	0.0145
	(0.0031)	(0.0920)	(0.0104)
Environmental	-0.0021	0.1180	0.0274*
	(0.0029)	(0.0884)	(0.0105)
Psychological	0.0026	-0.1420	-0.0289**
	(0.0025)	(0.0821)	(0.0092)
Social	0.0103***	0.3210***	0.0248**
	(0.0019)	(0.0715)	(0.0088)
Time	-0.0008	0.0486	0.0056
	(0.0025)	(0.0682)	(0.0088)
# years since last occupation	0.0046***	0.0183**	0.0028***
0 if in actual occupation	(0.0005)	(0.0057)	(0.0005)
Interactions: occupational demands × # of years			
Ergonomic × years	-0.0020**	0.0124	-0.0011
	(0.0007)	(0.0080)	(0.0006)
Environmental × years	0.0036***	-0.0013	0.0013*
	(0.0007)	(0.0058)	(0.0005)
Psychological × years	-0.0021**	-0.0102	-0.0001
	(0.0007)	(0.0065)	(0.0005)
Social × years	-0.0011*	0.0095*	-0.0017**
	(0.0005)	(0.0042)	(0.0005)
Time × years	0.0018**	-0.0064	0.0010*
	(0.0006)	(0.0060)	(0.0005)
F (all demands)	12.02***	20.47***	26.77***
Adj. R	0.012	0.102	0.064
# clusters	86	86	86
N	292,747	264,965	289,288

Note: All models refer to Equation 4.4 and include gender, age, $(\text{age}/10)^2$, marital status, West Germany and working hours/week as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

vidual's probability to smoke by 2.9 percentage points. In line with the previous results, we find a positive relation between the years since someone's last job and the probability to smoke. For the no longer employed the data show that higher environmental demands in the last job are correlated with the smoking probability. In contrast, the relationship between social demands and the probability to smoke decreases with each additional year out of the labor force. The F-test on joint significance again indicates that the occupational demands as a whole are significantly related to the smoking.

When we include the occupational demands one at a time (not shown), all occupational demands turn out to be significantly related to all three health outcomes, except for time demands. Taken as a whole, we conclude that also the second requirement for the mediation analyses is met.

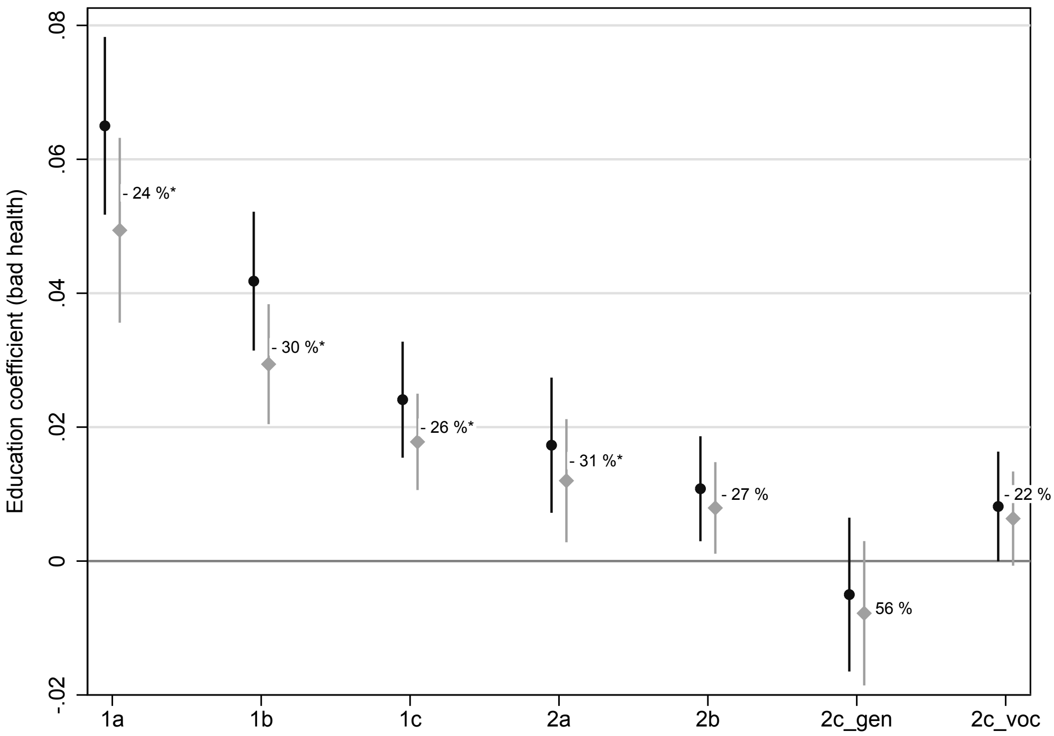
4.5.2 Mediating Role of Occupational Demands

As shown in the previous subsection, the requirements for the mediation analysis are generally met. That is, there is a path from education to occupational demands on the one hand and a path from occupational demands to health (behavior) on the other hand. Hence, we finally analyze to what extent the different occupational demands mediate the education-health and education-health-behavior relationships (Equation 4.5). Table 4.4 reports the results of both, Equation (4.2) and Equation (4.5) for the three health outcomes. To illustrate, we present these education coefficients graphically in Figures 4.3, 4.4 and 4.5. The gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % confidence intervals. For comparison, the education coefficients of the baseline model (Equation 4.2) and the corresponding 95 % confidence intervals are presented by the black circles. These coefficients again have to be interpreted in reference to tertiary education. The changes in the education coefficients across the two models expressed in percentage terms are presented within the graph and in Table 4.A.4 in the Appendix.

Bad Health

In line with previous studies, we find that education is in general significantly related bad health within the baseline model (black circles, Figure 4.3). We find that the five lowest educational levels are significantly and negatively related to bad health. This implies that compared to tertiary educated individuals, individuals with lower levels of education, have a higher probability of reporting bad health. For instance, primary-educated individuals have a 7 percentage-point higher probability of being in bad health compared to individuals with tertiary education. There seems to be no significant difference in health status between tertiary educated and upper secondary educated individuals. If at all, we expect occupational demands to mediate the relationship between the educational levels beyond upper secondary education and the probability to be in bad health.[72]

Figure 4.3: Change in education coefficients: Bad health



Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; The figures present the corresponding change in percentages (Equation 4.6); * $p<0.05$ (Table 4.A.4); *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Focusing on the differences between the education coefficients across the two models, we find that the estimates of the adjusted model (gray diamonds) are smaller compared to the estimates from the baseline model (black circles). The education estimates for primary (1a) up to intermediate secondary general education (2a) reduce significantly between 24 % and 31 % when all occupational demands are included simultaneously.[73] The test on joint significance of all educational dummies remains significant (see Table 4.4). While occupational demands seem to account for some of the differences in bad health between the lower up to intermediate educational groups and tertiary educated individuals, they do hardly account for health differences between upper secondary educated individuals and university graduates. Regarding the coefficients on occupational demands, only social

[72]There could still be mediations in the coefficients above upper secondary education but in unexpected directions.

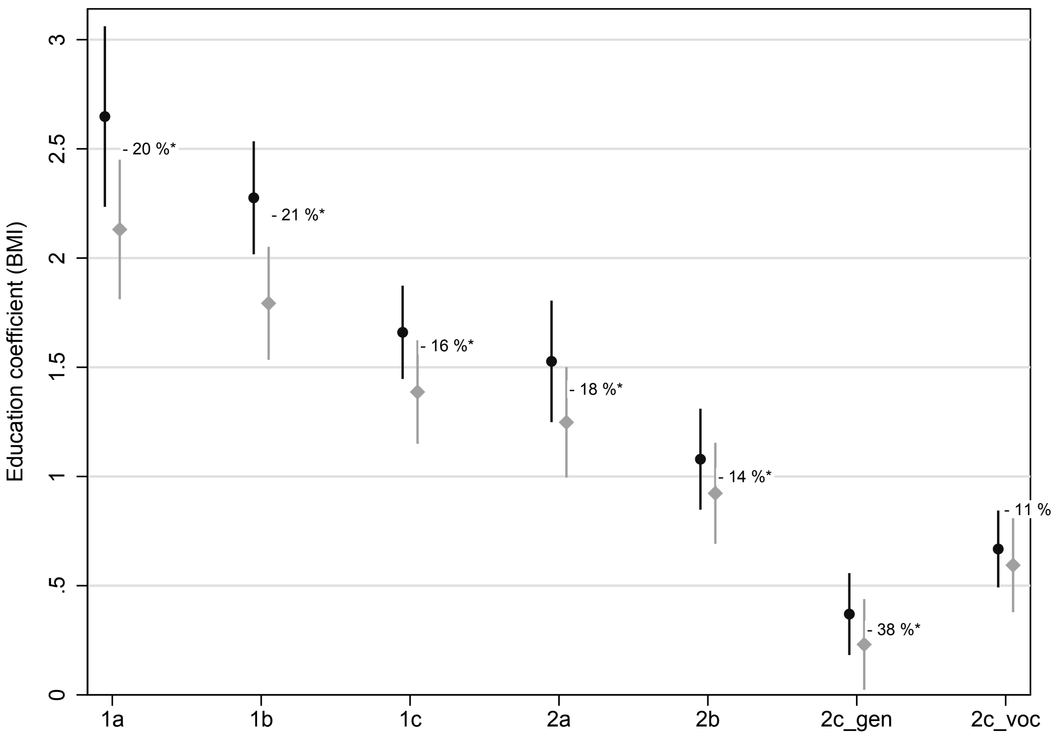
[73]It has to be considered that large changes in percentage terms also result from the very small estimates.

demands are significantly related to the probability of being in bad health. Interestingly, for the no longer employed individuals in our sample, we actually do find a significant relationship between the occupational demands in their last job and their health status. We further explore this in the robustness analyses.

BMI

Presented in Figure 4.4, we find that BMI is significantly related to all educational levels in comparison to tertiary education. The estimates gradually decline with increasing educational level stating an educational gradient up to upper secondary general education (2c_gen). The strongest relationship appears for primary educated individuals having a 2.6 kg/m^2 higher BMI than tertiary-educated individuals (reference group).

Figure 4.4: Change in education coefficients: BMI



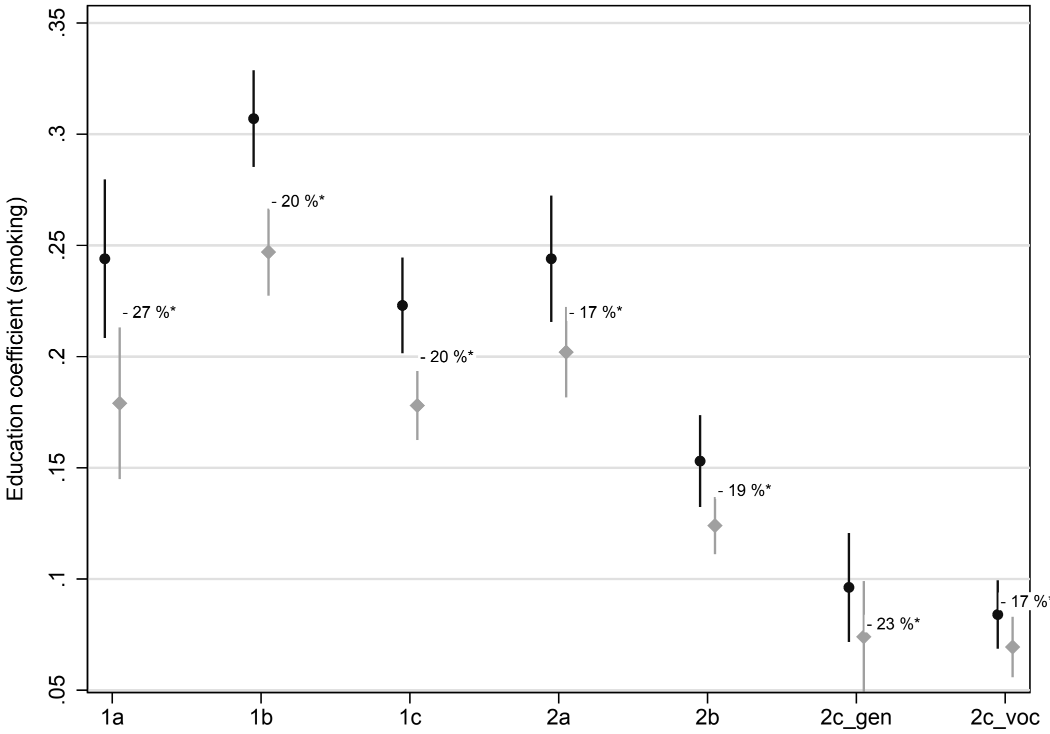
Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; The figures present the corresponding change in percentages (Equation 4.6); * $p<0.05$ (Table 4.A.4); *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Contrary to bad health, it appears that BMI is significantly higher among individuals with all other educational backgrounds, compared to tertiary education. Even after including occupational demands, all education coefficients as well as education as a whole (see joint F-test, Table 4.4) remain significantly related to BMI. Except for the upper secondary vocational educational level, all individual education coefficients reduce significantly in the range of 14–38 % when the occupational demands are included. These findings are comparable to previous research testing different mechanisms. For instance, Cutler and Lleras-Muney (2010) find that income accounts for 16 % of the education gradient in BMI, cognitive ability for 14 % and social integration for 18 %. Detailed analyses adding the occupational demands one at a time reveal that these reductions are mainly attributable to social demands within the occupation (see Figure 4.A.6 in the Appendix).

Smoking

The probability to smoke is also significantly related to education as a whole and across all educational levels (Figure 4.5 and Table 4.4). Individuals in all lower educational levels are significantly more likely to smoke compared to tertiary-educated individuals. This is similar to the finding for BMI, although there is no such definite educational gradient in smoking behavior. The highest probability of being a smoker can be found for individuals with lower secondary general education (1b), having a 31 percentage point higher probability to smoke than their tertiary-educated counterparts. But all educational estimates remain again significant at the 0.1 % level when all occupational demands are included (gray diamonds). Additionally, the F-test on joint significance of all educational dummies also remains significant (Table 4.4). Nevertheless, the magnitude of all education coefficients decreases significantly after adjusting for occupational demands. For instance, the change in the coefficient is most profound for primary education (1a) reducing by 27 % (from 0.244 to 0.179). Detailed analyses exploring the role of single occupational demands again suggest social demands but also ergonomic and environmental demands to be crucial (see Figure 4.A.7 in the Appendix). The extent to which all considered occupational demands partly mediate the educational differences in smoking is in a similar range as previously found for income (26 %) or other economic resources (33 %) (Cutler and Lleras-Muney, 2010). However, cognitive ability accounted for a larger share of about 45 %.

Figure 4.5: Change in education coefficients: Smoking



Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; The figures present the corresponding change in percentages (Equation 4.6); * $p<0.05$ (cf. Table 4.A.4); *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Table 4.4: OLS estimates: Mediation analyses (cf. Figures 4.3, 4.4, 4.5)

	Bad health		BMI		Smoker	
Model (Equation)	(1)	(4)	(1)	(4)	(1)	(4)
Education						
Primary (1a)	0.0650***	0.0494***	2.6480***	2.1310***	0.2440***	0.1789***
	(0.0068)	(0.0070)	(0.2110)	(0.1630)	(0.0182)	(0.0174)
Lower secondary gen (1b)	0.0418***	0.0294***	2.2760***	1.7930***	0.3068***	0.2466***
	(0.0053)	(0.0046)	(0.1320)	(0.1320)	(0.0111)	(0.0100)
Lower secondary voc (1c)	0.0241***	0.0178***	1.6600***	1.3870***	0.2231***	0.1782***
	(0.0044)	(0.0037)	(0.1090)	(0.1210)	(0.0110)	(0.0079)
Intermediate secondary gen (2a)	0.0173**	0.0120*	1.5270***	1.2479***	0.2442***	0.2024***
	(0.0052)	(0.0047)	(0.1420)	(0.1289)	(0.0145)	(0.0104)
Intermediate secondary voc (2b)	0.0108**	0.0079*	1.0790***	0.9230***	0.1530***	0.1241***
	(0.0040)	(0.0035)	(0.1180)	(0.1176)	(0.0105)	(0.0066)
Upper secondary gen (2c_gen)	-0.0050	-0.0078	0.3700***	0.2310*	0.0962***	0.0740***
	(0.0059)	(0.0055)	(0.0957)	(0.1055)	(0.0125)	(0.0128)
Upper secondary voc (2c_voc)	0.0082	0.0063	0.6680***	0.5940***	0.0840***	0.0694***
	(0.0042)	(0.0036)	(0.0898)	(0.1100)	(0.0078)	(0.0069)
Tertiary	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*
Occupational demands						
Ergonomic demands		0.0030		0.0144		0.0081
		(0.0031)		(0.0862)		(0.0074)
Environmental demands		-0.0037		-0.0195		0.0092
		(0.0031)		(0.0712)		(0.0071)
Psychological demands		0.0042		-0.0201		-0.0123*
		(0.0025)		(0.0556)		(0.0062)
Social demands		0.0086***		0.2471***		0.0160
		(0.0019)		(0.0649)		(0.0082)
Time demands		-0.0004		0.0757		0.0086
		(0.0024)		(0.0615)		(0.0073)
# years since last occupation	0.0032***	0.0045***	0.0150***	0.0064	0.0002	0.0013**
0 if in actual occupation	(0.0005)	(0.0005)	(0.0036)	(0.0044)	(0.0004)	(0.0004)
Interactions: occupational demands × # of years						
Ergonomic × years		-0.0020**		0.0153*		-0.0007
		(0.0007)		(0.0071)		(0.0005)
Environmental × years		0.0035***		-0.0010		0.0015***
		(0.0007)		(0.0051)		(0.0004)
Psychol. × years		-0.0020**		-0.0083		0.0001
		(0.0007)		(0.0060)		(0.0005)
Social × years		-0.00116*		0.0095*		-0.0017***
		(0.000452)		(0.0041)		(0.0005)
Time × years		0.00176**		-0.0085		0.0007
		(0.000581)		(0.0054)		(0.0004)
F (all education levels)	30.73***	15.92***	62.07***	54.79***	137.14***	130.77***
F (all demands)		7.91***		4.69***		15.49***
Adj. R	0.010	0.013	0.108	0.112	0.075	0.080
# clusters	86	86	86	86	86	86
N	292,352	292,352	264,634	264,634	288,898	288,898

Note: Model (1) refers to Equation (4.2), Model (4) to Equation (4.5); All models include gender, age, $(age/10)^2$, marital status, West Germany and working hours/week as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

4.5.3 Robustness Analyses

We perform two robustness checks. First, we exclude the non-working population to see whether our findings are the result of including those who no longer work. Second, we analyze the role of occupational demands in explaining the relation between education and self-reported health within one dataset. This is possible using the Employment Survey 2005/2006 solely as the data include a measure for self-reported health. The purpose of this last robustness check is to examine whether the moderate mediation results for our health status variable are due to the specific health measure within the German Microcensus 2009. Moreover, these analyses explore whether our results are potentially impaired by our merging procedure and possible differences across the two data sources.

Working Sample Only

In the main analyses we pool the working and non-working population as we consider it important to keep track of individuals who might have selected out of the labor force for health reasons. For the non-working individuals we take the occupational demands related to their last occupation interacted with the number of years out of the labor force. However, because our health measures rather capture current health (behavior) than the (long-term) stock of health (BMI might be seen as an exception), we also perform the mediation analyses for the working population solely.[74]

Table 4.5 shows that for health behavior (BMI and smoking) the results are very similar to the main results. Both the baseline relation between education and health behavior and the mediating effects very similar to the main findings. Given the comparable results we conclude that the findings for health behavior are unlikely driven by those who no longer work. This firstly implies that the relation between education and health behavior is independent from the working status, i.e. being employed is not endogenous to health behavior. Secondly, this implies that occupational demands do not seem to have a delayed effect on health behavior for the main part.

This observation does not hold for the correlation between education and bad health, which is much smaller in the working sample than in the full sample. Moreover, the relation between social demands and bad health is no longer significant. Both education and social demands are much stronger related to bad health for those who no longer work than for the working individuals (see Table 4.A.5 in the Appendix). This suggests that bad health might only in the long-run be related to social demands at work. An alternative explanation could be that those facing high social demands have a higher probability to leave the workforce. Research on longitudinal data should be performed to explicitly test this. Taken as a whole, the results regarding health status thus point to the fact

[74]In these robustness analyses, we also control for five dummies of company size which is not possible in the analyses in which we include the full sample.

Table 4.5: Robustness I: OLS estimates for the employed sample

	Bad health		BMI		Smoker	
Model (Equation)	(1)	(4)	(1)	(4)	(1)	(4)
Educational level						
Primary (1a)	0.0236**	0.0182*	2.400***	1.9721***	0.2282***	0.1593***
	(0.0071)	(0.0072)	(0.176)	(0.1541)	(0.0204)	(0.0211)
Lower secondary gen (1b)	0.0178**	0.0127*	2.180***	1.7767***	0.3117***	0.2457***
	(0.0054)	(0.0049)	(0.1210)	(0.1171)	(0.0123)	(0.0113)
Lower secondary voc (1c)	0.0119**	0.0083*	1.6121***	1.3528***	0.2258***	0.1793***
	(0.0041)	(0.0038)	(0.0999)	(0.1061)	(0.0112)	(0.0081)
Intermediate secondary gen (2a)	0.0105	0.0071	1.5394***	1.2534***	0.2383***	0.1923***
	(0.0055)	(0.0059)	(0.1381)	(0.1213)	(0.0146)	(0.0121)
Intermediate secondary voc (2b)	0.0058	0.0037	1.0838***	0.9217***	0.1541***	0.1254***
	(0.0040)	(0.0037)	(0.1012)	(0.0986)	(0.0107)	(0.0069)
Upper secondary gen (2c_gen)	0.0101	-0.0119	0.4051***	0.2600*	0.1049***	0.0832***
	(0.0063)	(0.0061)	(0.0841)	(0.1011)	(0.0118)	(0.0123)
Upper secondary voc (2c_voc)	0.00451	0.0036	0.7081***	0.6282***	0.0875***	0.0731***
	(0.0037)	(0.0036)	(0.0784)	(0.0966)	(0.0077)	(0.0071)
Tertiary (3a,b)	*reference*	*reference*	*reference*	*reference*	*reference*	*reference*
Occupational demands						
Ergonomic demands		0.0001		0.0241		0.0031
		(0.0034)		(0.0881)		(0.0070)
Environmental demands		0.0005		0.0367		0.0112
		(0.0034)		(0.0688)		(0.0072)
Psychological demands		-0.0012		0.0589		-0.0113
		(0.0022)		(0.0525)		(0.0064)
Social demands		0.0016		0.2156**		0.0206*
		(0.0020)		(0.0676)		(0.0089)
Time demands		0.0028		0.0773		0.0090
		(0.0024)		(0.0565)		(0.0070)
F (all education levels)	6.69***	4.55***	67.68***	55.90***	121.05***	111.57***
F (all demands)		1.19		3.89**		19.70***
Adj. R	0.003	0.004	0.121	0.123	0.063	0.067
# clusters	86	86	86	86	86	86
N	224,683	224,683	202,662	202,662	221,437	221,437

Note: Model (1) refers to Equation (4.2), Model (4) to Equation (4.5); All models include gender, age, $(age/10)^2$, marital status, West Germany, working hours/week and dummies for the company size as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

that lifetime exposure is more relevant for health status compared to current exposure, which has been previously shown (Monden, 2005; Fletcher et al., 2011). In contrast to our measures of health behavior, the robustness analyses suggest that employment is likely endogenous regarding health status.

Employment Survey 2005/2006 Data Only

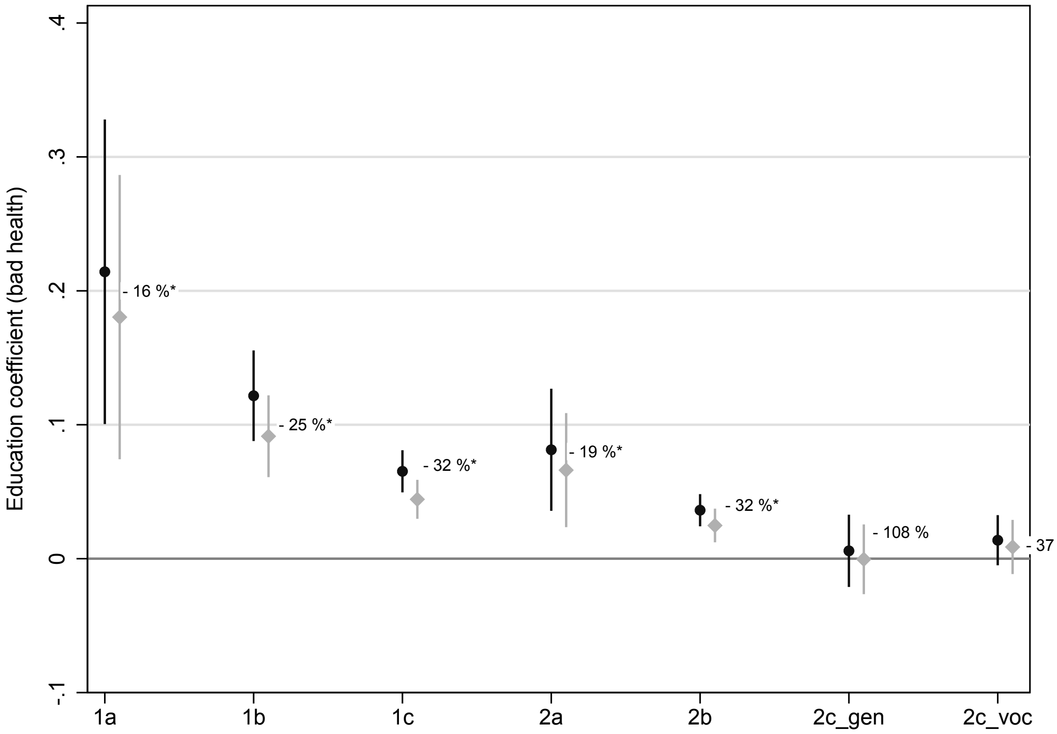
The German Employment Survey 2005/2006 includes self-reported health of workers on a five-point Likert scale. In contrast to the health status measure in the Microsensus (sickness in past four weeks), self-reported health likely takes into account chronic and acute illnesses (e.g. Simon et al., 2005) and has been found to be a predictor for mortality (e.g. Idler and Benyamini, 1997). We define individuals who reported their health status to be “fair” or “poor” to be in bad health. Figure 4.6 illustrates the results of the mediation analysis. It has to be considered that the Employment Survey only includes employed individuals. The relation between education and bad self-reported health (Employment Survey) is much stronger than the relation between education and bad health (Microcensus), i.e. the education estimates are much larger in Figure 4.6 compared to those presented in Table 4.5. The specific nature of the health measure in the Microcensus data, which rather captures the current short-term health status of the respondents, might be responsible for the different results. However, the mediating role of occupational demands is still moderate and comparable to those in the previous analyses. This suggests that also for other measures of subjective health, occupational demands only account for educational differences in health for the lower educational groups.

Our results indicate that occupational demands mediate the relationship between education and health less strongly compared to studies from other countries focusing on the working population (e.g. Warren et al., 2004; Brand et al., 2007; Niedhammer et al., 2008). The differences might be explained by the different measures of occupational demands. Previous studies include self-reported and thus subjective occupational demands, which might capture personal characteristics to some extent. Given the specific construction of the demands in this study, we claim our measures to be more objective.

4.6 Conclusions

The aim of the study was to analyze the potential mediating role of occupational demands in the education-health (behavior) relationship. Our first contribution is that we focus on the role of five specific occupational demands: ergonomic, environmental, psychological, social and time demands. Previous studies most often either include occupations by occupational dummies or only differentiate between physical and psychosocial demands. Both approaches disregard mapping the work environment of occupations in a comprehensive way. This is of particular importance as our results indicate that education is

Figure 4.6: Robustness II: OLS estimates based on the Employment Survey



Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; The figures present the corresponding change in percentages (Equation 4.6); * $p<0.05$ (Table 4.A.6); *Source:* Employment Survey 2005/2006.

related to psychological and social occupational demands in different directions. Our second contribution is that we do not solely look into the mediating role of occupational demands between education and subjective health status, but also into the mediating role of occupational demands in the relationship between education and health behavior. In this context, we analyze BMI and smoking. Our third contribution is that we extend the analyses by focusing on a broad sample. We include the whole working population instead of focusing on a subgroup, e.g. high educated workers. We additionally include individuals who do (currently) not actively participate in the labor market.

First, the results indicate that the lower educated have on average higher and more hazardous ergonomic, environmental and social occupational demands compared to individuals with a higher educational qualification. In contrast, higher educated individuals tend to suffer from higher psychological demands more often. Time demands are not significantly related to education. When all demands are included simultaneously, especially social demands turn out to be significantly related to the probability of being in bad health, to the probability to smoke and to BMI. Second, we find that occupational demands partially mediate the relationship between education and health behavior. Education coefficients on BMI and smoking significantly reduce up to 38 % when the occupational demands are included. Especially social demands seem to play a major role. The extent to which the occupational demands considered mediate the educational differences in BMI and smoking is comparable to other factors that have been previously found to mediate the education gradient in health behavior, like income, economic resources or social integration (Cutler and Lleras-Muney, 2010). Regarding our health status measure, the results suggest that occupational demands account for the differences in bad health between lower and intermediately educated individuals on the one hand and tertiary educated individuals on the other up to 30 %. Especially for low and intermediate educated individuals. However, they do hardly account for health differences between individuals with upper secondary vocational education and university graduates. These results suggest that taking a differentiated measure of education is crucial for such mediation analyses. Our results regarding our measure of subjective health status deviate from previous studies which focused on subpopulations, such as high-school graduates only (e.g. Warren et al., 2004; Brand et al., 2007). They found that socioeconomic differences in subjective health status were reduced to a greater extent when controlling for physical and psychosocial work factors.

This study reveals several opportunities for future research. First, while previous studies have only focused on educational differences in health status, our study shows that occupational demands also account for differences in health behavior. Future research should thus also focus on other preceding behavioral outcomes, such as health utilization or health care spending to improve our understanding. Second, it seems important to take the entire working population and especially not employed individuals into account, when analyzing the mediating factors of the relationship between education and health status. In contrast to health behavior, our findings indicate that past occupational demands are more relevant for the respondent's health than current exposure. Future research should focus on working histories and take the specific occupational (and more objective) occupational demands to explicitly test this. Relatedly, there is a need to analyze the role of specific occupational demands in mediating the educational differences in the outflow of (older) workers out of the active labor market into early retirement and long-term sickness. This is an important issue also for policy makers in the context of raising the retirement age in most European countries.

In conclusion, this study provides suggestive evidence that there might be important dynamic effects in explaining the relation between education and health (behavior) via work-related conditions, such as occupational demands. Moreover, our findings indicate that existing inequalities in the working conditions do matter for the educational differences in BMI and smoking. Improving the working conditions especially for lower educated individuals might thus contribute to a reduction of educational differences in health.

Appendix

Table 4.A.1: Summary statistics

Variables	Employment Survey 2006 Mean (SD)	German Microcensus 2009 Mean (SD)
Controls		
Female	0.484	0.498
Age	42.13 (9.44)	45.49 (10.91)
# of working hours	39.37 (13.17)	36.60 (10.06)[a]
Married	0.547	0.633
West Germany	0.811	0.792
Employed	1	0.773
Education (CASMIN)		
Primary (1a)	0.005	0.024
Lower secondary gen (1b)	0.027	0.070
Lower secondary voc (1c)	0.202	0.258
Intermediate secondary gen (2a)	0.017	0.024
Intermediate secondary voc (2b)	0.316	0.324
Upper secondary gen (2c_gen)	0.019	0.021
Upper secondary voc (2c_voc)	0.120	0.099
Tertiary (3a,b)	0.294	0.181
Main occupations (KldB 1992)		
Forestry, agriculture	0.018	0.025
Mining	0.001	0.002
Manufacturing	0.185	0.211
Engineering	0.082	0.063
Services	0.711	0.688
Others	0.004	0.011
Health outcomes		
Bad health	0.095[b]	0.127
BMI	n.a.	25.66 (4.45)
Smoker	n.a.	0.317
N	18,797	359,587

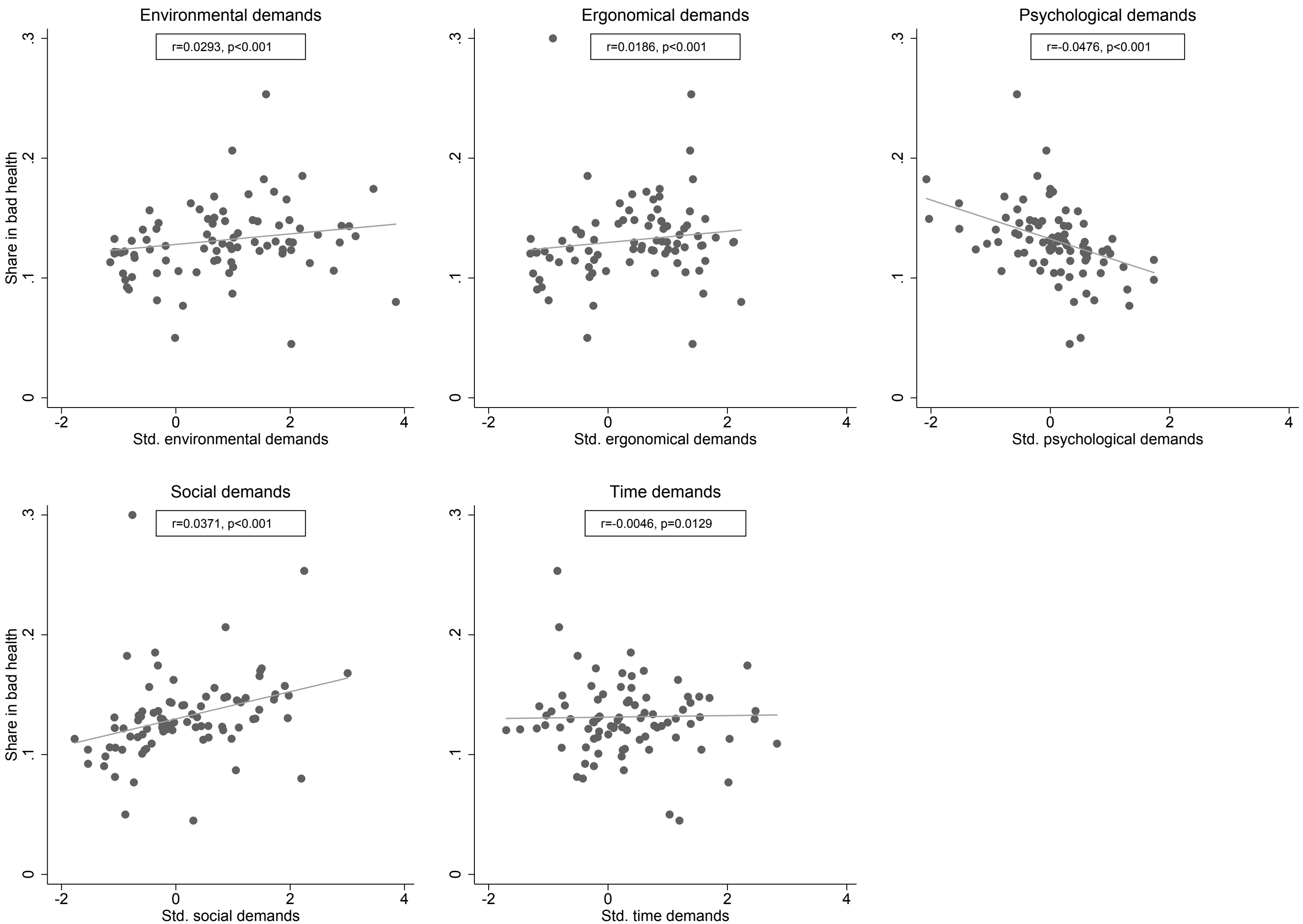
Note: [a]Number of working hours for employed individuals, 0 if currently not employed; [b]Based on self rated health scale within the Employment Survey; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

Table 4.A.2: Summary statistics of occupational demand items

Items: classification cf. Kroll (2011)	Mean	(SD)
Ergonomic demands		
Working in a standing position	0.510	(0.500)
Sedentary work	0.112	(0.315)
Lift and carry heavy charges	0.198	(0.398)
Working in forced positions	0.121	(0.327)
Environmental demands		
Working in dust, gas, fumes	0.117	(0.322)
Working in cold, heat, under wet conditions	0.181	(0.385)
Working in dirt, oil, grease	0.144	(0.351)
Working with vibrations	0.038	(0.192)
Working with glaring/faint light	0.087	(0.282)
Working with dangerous substances	0.065	(0.246)
Wearing protection clothes while work	0.181	(0.385)
Working under noise	0.210	(0.408)
Working while others smoke	0.153	(0.360)
Psychological demands		
Deadline pressure/pressure to perform	0.566	(0.496)
Getting new tasks and become acquainted with it	0.428	(0.495)
Trying new or improve procedures	0.306	(0.461)
Being interrupted, disturbed while working	0.498	(0.500)
Minimum output or time is prescribed	0.303	(0.460)
Doing unlearned tasks	0.093	(0.291)
Doing or observing different tasks simultaneously	0.630	(0.483)
Minor mistakes have major financial consequences	0.157	(0.364)
Stretching to the limits of own performance/abilities	0.172	(0.377)
Working very fast	0.439	(0.496)
Social demands		
Feeling as a part of community	0.031	(0.172)
Good cooperation with colleagues	0.006	(0.077)
Being supported by colleagues	0.018	(0.132)
Being supported by direct superior	0.055	(0.227)
Arranging work on own schedule	0.059	(0.236)
Having influence on amount of work	0.238	(0.426)
Allotting breaks autonomously	0.183	(0.387)
Feeling that own work is important	0.146	(0.120)
Not being on notice of changes, decisions	0.143	(0.350)
Being poorly informed about own work	0.087	(0.281)
Time demands		
Working in shifts	0.226	(0.418)
Working on call	0.200	(0.400)
Working sometimes on Saturdays	0.682	(0.466)
Working sometimes on Sundays and holidays	0.443	(0.497)
Working sometimes between 11pm an 5am	0.230	(0.421)

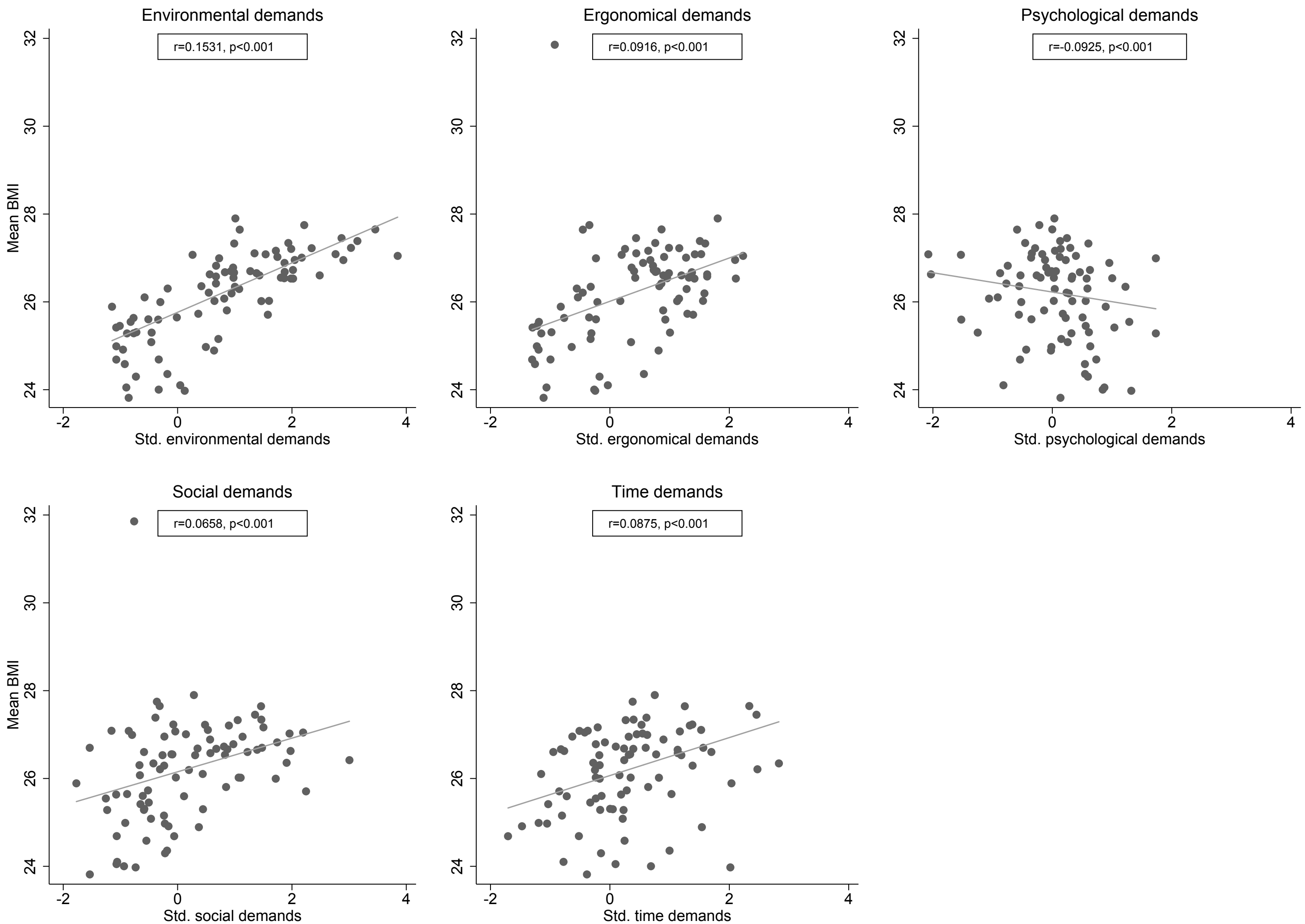
Source: Employment Survey 2005/2006.

Figure 4.A.1: Occupational demands and bad health



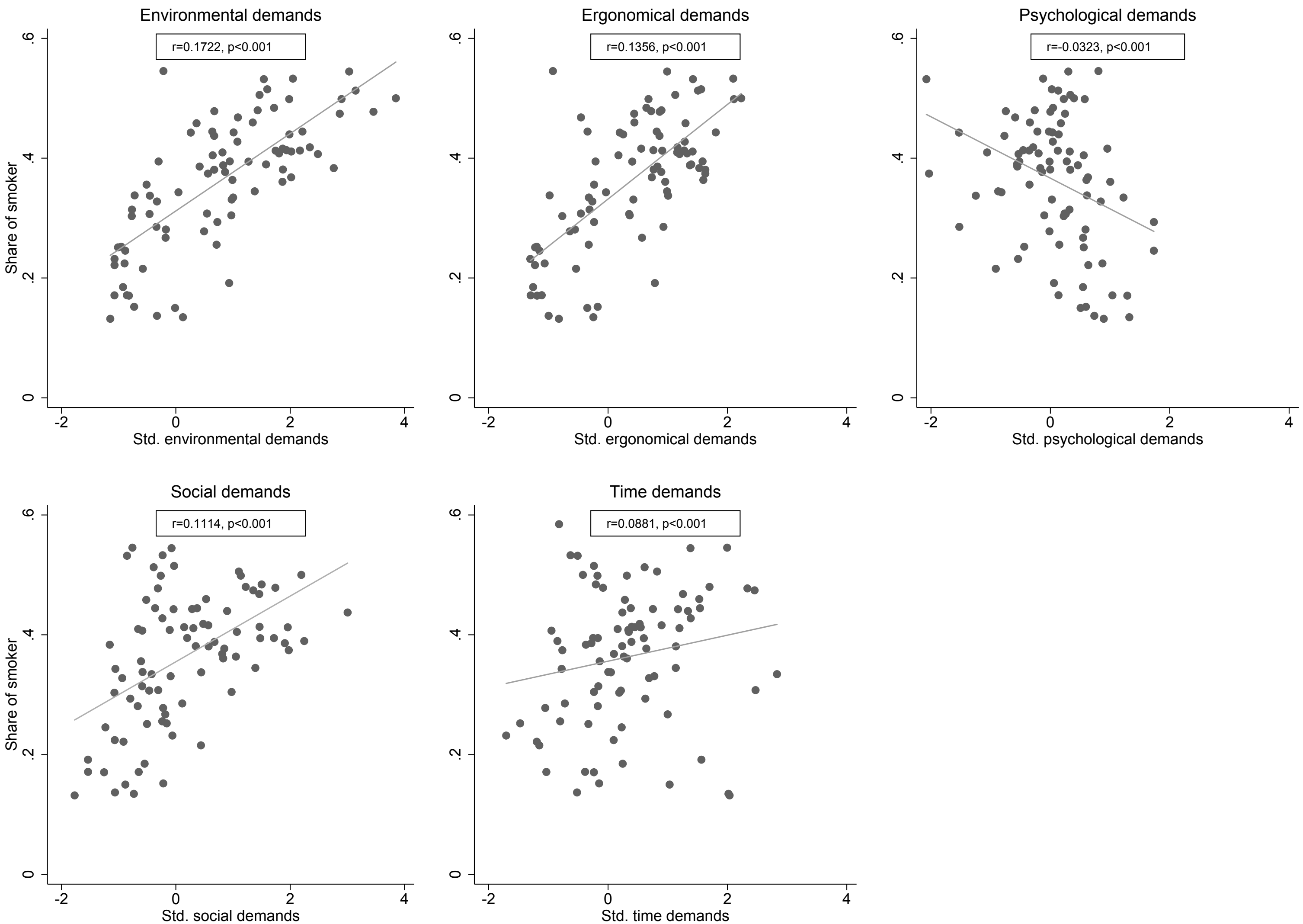
Note: Average share of individuals in bad health by occupational demands within occupations; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Figure 4.A.2: Occupational demands and BMI



Note: Average BMI by occupational demands within occupations; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Figure 4.A.3: Occupational demands and smoking



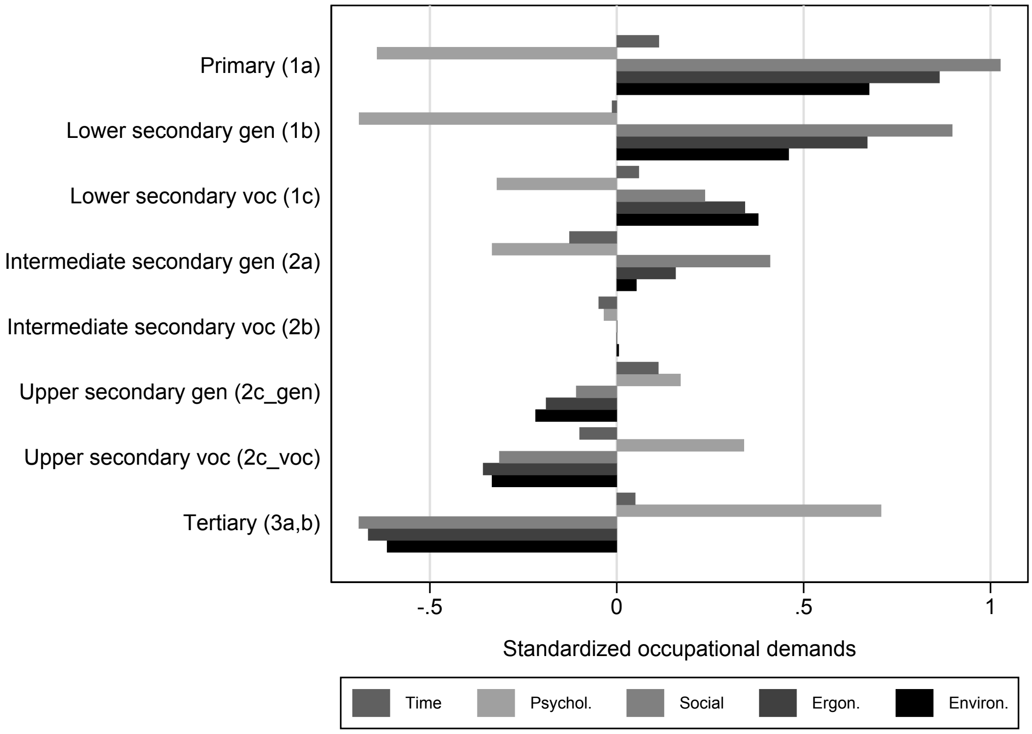
Note: Average share of smoker by occupational demands within occupations; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Table 4.A.3: Correlation matrix of occupational demands (aggregated by occupations)

	Ergonomic.	Environ.	Psychol.	Social	Time
Ergonomic	1				
Environmental	0.721***	1			
Psychological	-0.365***	-0.127	1		
Social	0.466***	0.416***	-0.320**	1	
Time	0.096	0.293**	0.226*	-0.017	1

Note: * p<0.05, ** p<0.01, *** p<0.001; *Source:* Employment Survey 2005/2006.

Figure 4.A.4: Occupational demands by educational level (CASMIN)



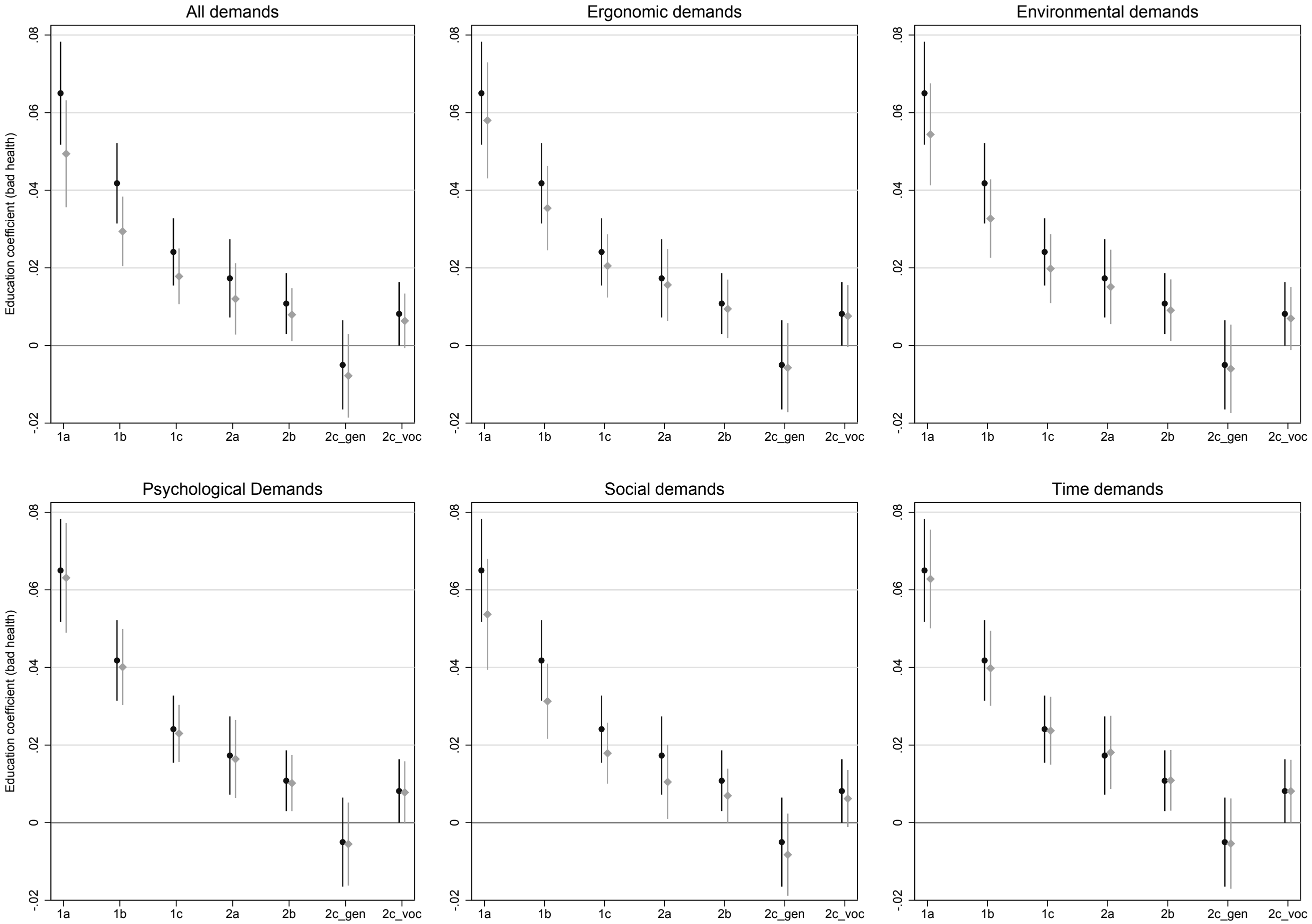
Note: The different bars show the mean occupational demands across all occupations within one educational level; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Table 4.A.4: Change in coefficients in % (cf. Figures 4.3, 4.4, 4.5; Tab 4.4)

	Bad health	BMI	Smoker
Primary (1a)	-24.00*	-19.52*	-26.64*
	(3.92)	(3.31)	(6.96)
Lower secondary gen (1b)	-29.67*	-21.22*	-19.54*
	(3.47)	(3.54)	(6.34)
Lower secondary voc (1c)	-26.14*	-16.45*	-20.18*
	(2.50)	(3.03)	(5.19)
Intermediate secondary gen (2a)	-30.64*	-18.27*	-17.21*
	(1.98)	(2.73)	(4.29)
Intermediate secondary voc (2b)	-26.57	-14.46*	-18.95*
	(1.71)	(2.33)	(3.66)
Upper secondary gen (2c_gen)	55.69	-37.57*	-23.08*
	(1.68)	(2.20)	(2.87)
Upper secondary voc (2c_voc)	-22.21	-11.08	-17.38*
	(1.56)	(1.50)	(2.46)
Tertiary	*Reference*	*Reference*	*Reference*

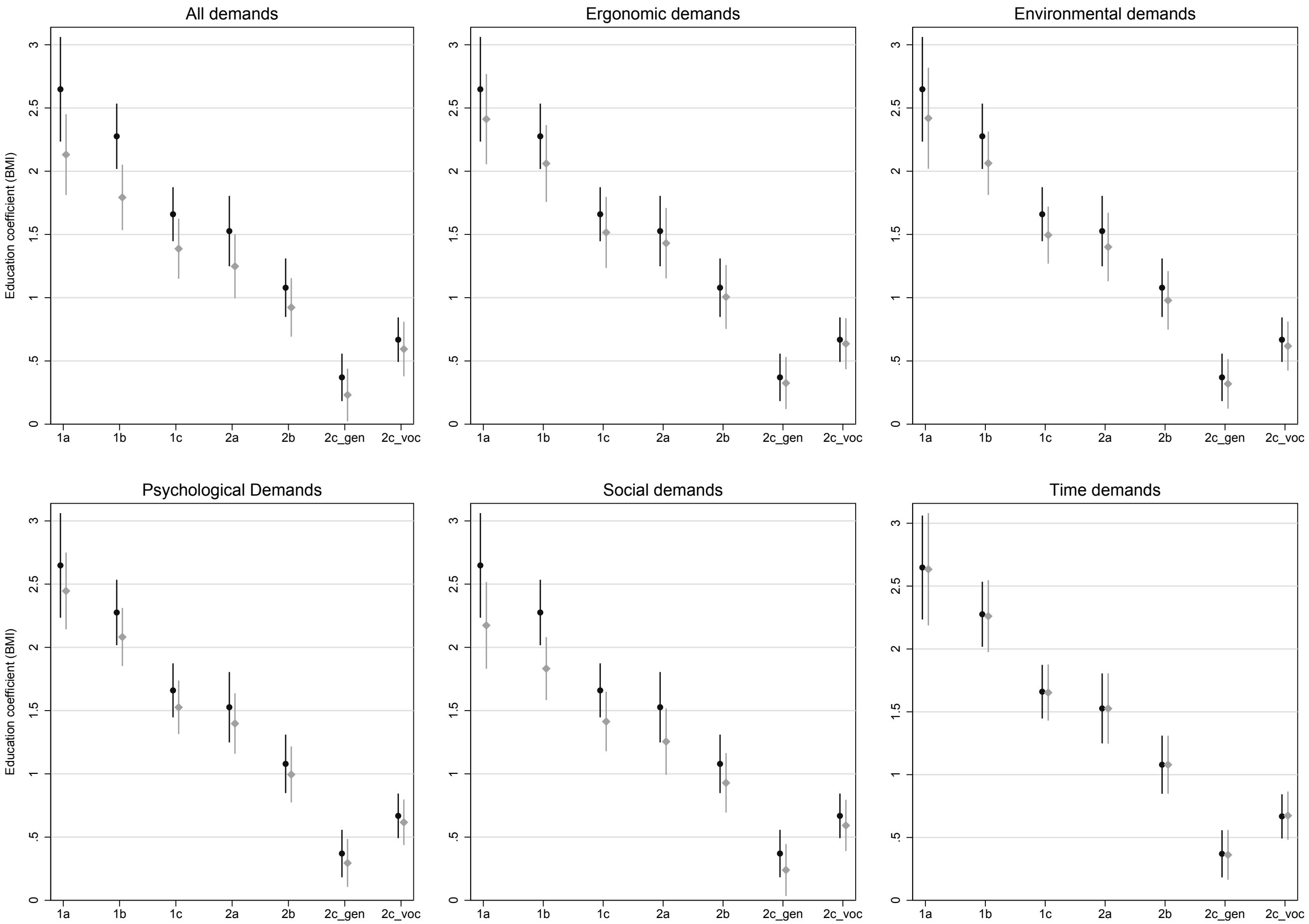
Note: $\%change = \frac{\theta_1 - \gamma_1}{\gamma_1} \times 100$ cf. Equations 4.2 and 4.5; Bootstrapped standard errors (500 reps), t-values in parentheses;* p<0.05; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

Figure 4.A.5: Change in education coefficients: bad health (more detailed)



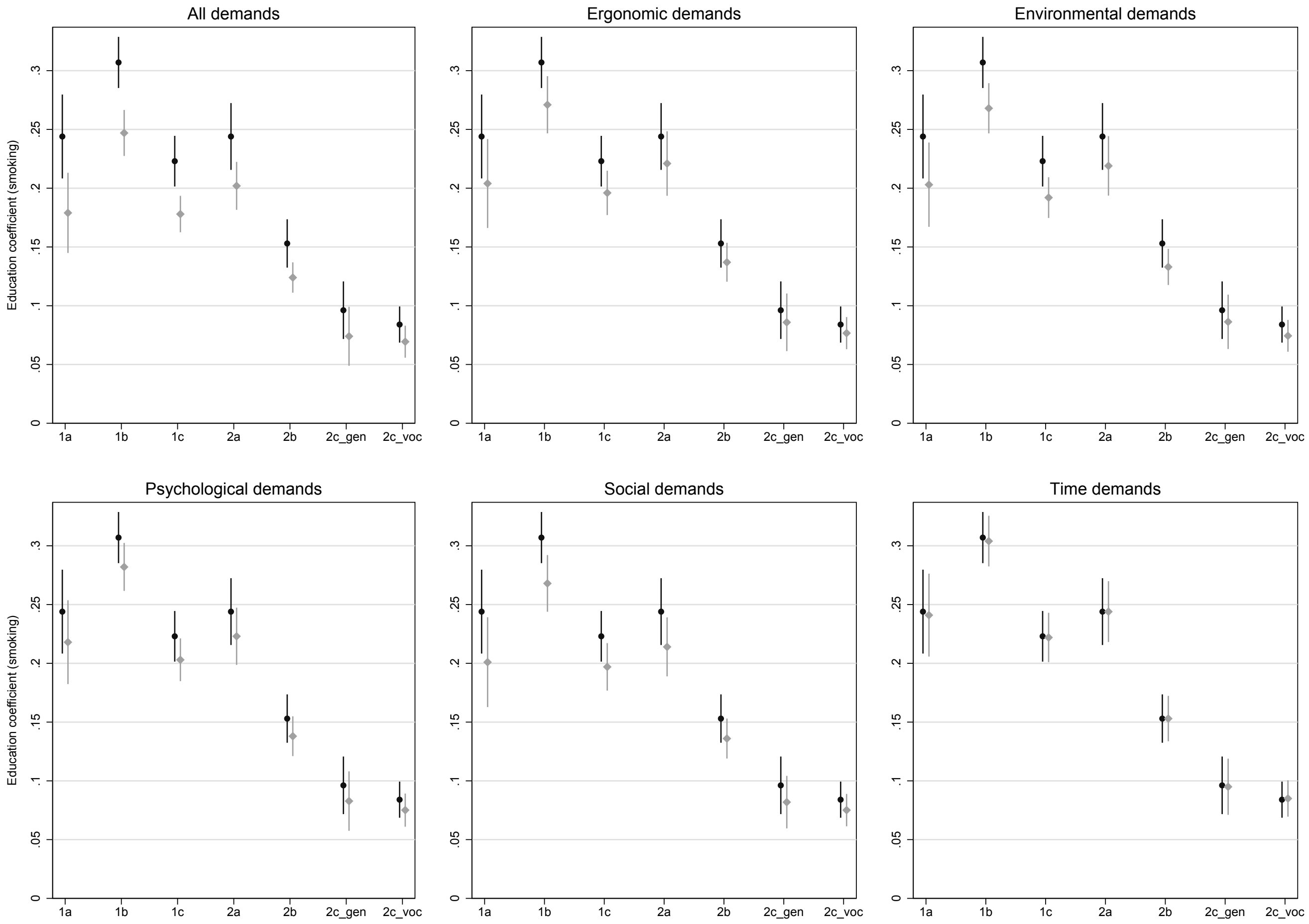
Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; Education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Figure 4.A.6: Change in education coefficients: BMI (more detailed)



Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; Education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Figure 4.A.7: Change in education coefficients: smoking (more detailed)



Note: Black circles belong to the baseline education coefficients (Equation 4.2) and the corresponding 95 % CI, gray diamonds belong to the education coefficients adjusted for all occupational demands (Equation 4.5) and the corresponding 95 % CI; In reference to tertiary education; Education: 1a primary, 1b lower secondary gen, 1c lower secondary voc, 2a intermediate secondary gen, 2b intermediate secondary voc, 2c_gen upper secondary gen, 2c_voc upper secondary voc; *Source:* German Microcensus 2009; Employment Survey 2005/2006.

Table 4.A.5: Robustness analyses Ib: OLS estimates for the unemployed sample

	Bad health		BMI		Smoker	
Model (Equation)	(1)	(4)	(1)	(4)	(1)	(4)
Educational level						
Primary (1a)	0.1080***	0.0929***	2.715***	2.228***	0.1923***	0.1612***
	(0.0123)	(0.0116)	-0.214	(0.244)	(0.0196)	(0.0186)
Lower secondary gen (1b)	0.0808***	0.0693***	2.205***	1.737***	0.2461***	0.2169***
	(0.0081)	(0.0074)	-0.18	(0.209)	(0.0122)	(0.0124)
Lower secondary voc (1c)	0.0609***	0.0565***	1.682***	1.438***	0.1747***	0.1532***
	(0.0071)	(0.0059)	-0.181	(0.197)	(0.0119)	(0.0105)
Intermediate secondary gen (2a)	0.0366***	0.0335***	1.333***	1.130***	0.2000***	0.1817***
	(0.0101)	(0.0090)	-0.213	(0.218)	(0.0142)	(0.0125)
Intermediate secondary voc (2b)	0.0344***	0.0339***	1.046***	0.931***	0.1152***	0.1022***
	(0.0064)	(0.0058)	-0.209	(0.210)	(0.0101)	(0.0089)
Upper secondary gen (2c_gen)	0.0310**	0.0284*	0.309	0.195	0.0191	0.0091
	(0.0112)	(0.0108)	-0.217	(0.215)	(0.0212)	(0.0217)
Upper secondary voc (2c_voc)	0.0254**	0.0255***	0.487**	0.435*	0.0388***	0.00327**
	(0.0085)	(0.0069)	-0.184	(0.204)	(0.0112)	(0.0112)
Tertiary (3a,b)	reference	reference	reference	reference	reference	reference
Occupational demands						
Ergonomic demands		0.0176***		0.0765		0.0163*
		(0.0048)		(0.1010)		(0.0066)
Environmental demands		-0.0168***		-0.0041		0.0041
		(0.0047)		(0.0864)		(0.0064)
Psychological demands		0.0185***		0.0721		-0.0074
		(0.0049)		(0.0931)		(0.0064)
Social demands		0.0164***		0.2551***		0.0064
		(0.0040)		(0.0735)		(0.0055)
Time demands		0.00141		0.1142		0.0091
		(0.0041)		(0.0870)		(0.0074)
# years since last occupation	-0.0002	0.0005	-0.0100*	-0.0204***	-0.0012***	-0.0001**
0 if in actual occupation	(0.0003)	(0.0003)	(0.0041)	(0.0051)	(0.0002)	(0.0003)
Interactions: occupational demands × # of years						
Ergonomic × years		-0.0020***		0.0061		-0.0005
		(0.0004)		(0.0066)		(0.0004)
Environmental × years		0.0026***		0.0079		0.0003
		(0.0005)		(0.0040)		(0.0003)
Psychol. × years		-0.0017***		-0.0069		-0.0003
		(0.0005)		(0.0054)		(0.0004)
Social × years		-0.0010**		0.0069*		-0.0005
		(0.0004)		(0.0029)		(0.0003)
Time × years		0.0011**		-0.0081		0.0001
		(0.0004)		(0.0055)		(0.0004)
F (all education levels)	23.52***	19.39***	43.35***	30.52***	88.30***	71.56***
F (all demands)		10.39***		5.41***		10.96***
adj. R	0.023	0.026	0.070	0.075	0.135	0.137
# clusters	86	86	86	86	86	86
N	67,689	67,689	61,990	61,990	67,480	67,480

Note: Model (1) refers to Equation (4.2), Model (4) to Equation (4.5); all models include gender, age, $(\text{age}/10)^2$, marital status, West Germany and working hours/week as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* German Microcensus 2009 and Employment Survey 2005/2006.

Table 4.A.6: Robustness analyses II: OLS estimates for bad health based on the Employment Survey 2005/2006

	(1) Bad health	(4) Bad health	(4a) Bad health
Educational level			
Primary (1a)	0.2142***	0.1804**	0.1707**
	(0.0580)	(0.0542)	(0.0553)
Lower secondary gen (1b)	0.1217***	0.0914***	0.0823***
	(0.0173)	(0.0156)	(0.0157)
Lower secondary voc (1c)	0.0652***	0.0443***	0.0351***
	(0.0080)	(0.0074)	(0.0071)
Intermediate secondary gen (2a)	0.0813***	0.0661**	0.0561*
	(0.0233)	(0.0217)	(0.0214)
Intermediate secondary voc (2b)	0.0362***	0.0248***	0.0177**
	(0.0061)	(0.0064)	(0.0062)
Upper secondary gen (2c_gen)	0.0058	-0.0005	-0.0066
	(0.0138)	(0.0133)	(0.0134)
Upper secondary voc (2c_voc)	0.0138	0.0087	0.0033
	(0.0096)	(0.0103)	(0.0102)
Tertiary (3a,b)	*reference*	*reference*	*reference*
Occupational demands			
Ergonomic demands		0.0109**	0.0101***
		(0.0038)	(0.0025)
Environmental demands		-0.0035	0.0202***
		(0.0037)	(0.0036)
Psychological demands		0.0102*	0.0213***
		(0.0041)	(0.0026)
Social demands		0.0188***	0.0336***
		(0.0036)	(0.0030)
Time demands		0.0024	-0.0035
		(0.0038)	(0.0027)
F (all education levels)	15.56***	12.67***	9.37***
F (all demands)		14.81***	70.88***
adj. R	0.028	0.032	0.059
# clusters	86	86	86
N	18,588	18,588	18,588

Note: Model (1) refers to Equation (4.2), Model (4) to Equation (4.5) with occupational demands generated via multilevel compared to the main analyses, Model (4a) to Equation (4.5) with occupational demands based on standardized individual sum scores; all models include gender, age, $(age/10)^2$, marital status, West Germany and working hours/ week as covariates; In reference to tertiary education; Clustered standard errors (occupation KldB 2-digit) in parentheses; * p<0.05, ** p<0.01, *** p<0.001; *Source:* Employment Survey 2005/2006.

Bibliography

Albertsen, K., Hannerz, H., Borg, V. and Burr, H. (2004). Work environment and smoking cessation over a five-year period, *Scandinavian Journal of Public Health* **32**(3): 164–171.

Anderson, P. M., Butcher, K. F. and Levine, P. B. (2003). Maternal employment and overweight children, *Journal of Health Economics* **22**(3): 477–504.

Anderson-Sprecher, R. (1994). Model comparisons and R-Squared, *The American Statistician* **48**(2): 113–117.

Andersson, M. A. and Maralani, V. (2015). Early-life characteristics and educational disparities in smoking, *Social Science & Medicine* **144**: 138–147.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors, *Journal of Business & Economic Statistics* **19**(1): 2–28.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, Princeton.

Aughinbaugh, A. and Gittleman, M. (2004). Maternal employment and adolescent risky behavior, *Journal of Health Economics* **23**(4): 815–838.

Bakker, A. B. and Demerouti, E. (2007). The job demands–resources model: state of the art, *Journal of Managerial Psychology* **22**(3): 309–328.

Barker, D. (1998). *Mothers, Babies and Health in Later Life*, Churchill Livingstone, Edinburgh.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *Journal of Personality and Social Psychology* **51**(6): 1173–1182.

Bauer, K. W., Hearst, M. O., Escoto, K., Berge, J. M. and Neumark-Sztainer, D. (2012). Parental employment and work-family stress: Associations with family food environments, *Social Science & Medicine* **75**(3): 496–504.

Becker, G. S. (2007). Health as human capital: synthesis and extensions, *Oxford Economic Papers* **59**(3): 379–410.

Becker, G. S. and Mulligan, C. B. (1997). The endogenous determination of time preference, *The Quarterly Journal of Economics* **112**(3): 729–758.

Biesanz, J. C., Falk, C. F. and Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects, *Multivariate Behavioral Research* **45**(4): 661–701.

Bishop, J. (2011). The effect of maternal employment on youth overweight in Australia, *Economic Record* **87**: 92–104.

Blane, D. (2006). The life course, the social gradient, and health, *in* M. Marmot and R. Wilkinson (eds), *Social Determinants of Health*, Oxford University Press, Oxford, pp. 54–77.

BMFSFJ (2012). Dossier Müttererwerbstätigkeit - Erwerbstätigkeit, Erwerbsumfang und Erwerbsvolumen 2012, *Report 2012*, Bundesministerium für Familie, Senioren, Frauen und Jugend, Wiesbaden.

Borg, V. and Kristensen, T. S. (2000). Social class and self-rated health: can the gradient be explained by differences in life style or work environment?, *Social Science & Medicine* **51**(7): 1019–1030.

Bound, J., Geronimus, A. T., Rodriguez, J. M. and Waidmann, T. (2014). The implications of differential trends in mortality for social security policy, *Research Paper 2014-314*, Michigan Retirement Research Center.

Brand, J. E., Warren, J. R., Carayon, P. and Hoonakker, P. (2007). Do job characteristics mediate the relationship between SES and health? Evidence from sibling models, *Social Science Research* **36**(1): 222–253.

Brauns, H., Scherer, S. and Steinmann, S. (2003). The CASMIN educational classification in international comparative research, *in* J. Hoffmeyer-Zlotnik and C. Wolf (eds), *Advances in Cross-National Comparison*, Springer US, pp. 221–244.

Bricard, D., Jusot, F., Beck, F., Khlat, M. and Legleye, S. (2015). Educational inequalities in smoking over the life cycle: an analysis by cohort and gender, *International Journal of Public Health* **61**(1): 101–109.

Bright, B. C. and Soulakova, J. N. (2014). Evidence of telescoping in regular smoking onset age, *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco* **16**(6): 717–724.

Brunello, G., Fort, M., Schneeweis, N. and Winter-Ebmer, R. (2016). The causal effect of education on health: What is the role of health behaviors?, *Health Economics* **25**(3): 314–336.

Bullock, J. G., Green, D. P. and Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer), *Journal of Personality and Social Psychology* **98**(4): 550–558.

Burkhauser, R. V. and Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research, *Journal of Health Economics* **27**(2): 519–529.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into american public schools, *The Journal of Human Resources* **44**(1): 140–170.

Case, A. C. and Deaton, A. (2003). Broken down by work and sex: how our health declines, *Working Paper Series 9821*, National Bureau of Economic Research.

Cawley, J. (2004). The impact of obesity on wages, *Journal of Human Resources* **XXXIX**(2): 451–474.

Cawley, J. (2010). The economics of childhood obesity, *Health Affairs (Project Hope)* **29**(3): 364–371.

Cawley, J. (2011). *The Oxford Handbook of the Social Science of Obesity*, Oxford University Press, New York.

Cawley, J. and Liu, F. (2012). Maternal employment and childhood obesity: A search for mechanisms in time use data, *Economics & Human Biology* **10**(4): 352–364.

Choo, E. and Denny, M. (2006). Wearing out – the decline in health, *Working Papers*, University of Toronto, Department of Economics.

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain, *The American Economic Review* **103**(6): 2087–2120.

Cole, T. J., Bellizzi, M. C., Flegal, K. M. and Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey, *BMJ (Clinical research ed.)* **320**(7244): 1240–1243.

Conti, G. and Hansman, C. (2012). Personality and the education–health gradient: A note on "understanding differences in health behaviors by education", *Journal of Health Economics* **32**(2): 480–485.

Conti, G. and Heckman, J. J. (2010). Understanding the early origins of the education–health gradient: A framework that can also be applied to analyze gene–environment interactions, *Perspectives on Psychological Science* **5**(5): 585–605.

Conti, G., Heckman, J. and Urzua, S. (2010). The education-health gradient, *The American Economic Review* **100**(2): 234–238.

Cropper, M. L. (1977). Health, investment in health, and occupational choice, *The Journal of Political Economy* **85**(6): 1273–1294.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development, *Journal of Economic Literature* **47**(1): 87–122.

Cutler, D. M. and Lleras-Muney, A. (2006). Education and health: Evaluating theories and evidence, *Working Paper 12352*, National Bureau of Economic Research.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education, *Journal of Health Economics* **29**(1): 1–28.

Datar, A., Nicosia, N. and Shier, V. (2014). Maternal work and children's diet, activity, and obesity, *Social Science & Medicine* **107**: 196–204.

de Walque, D. (2007). Does education affect smoking behaviors? evidence using the vietnam draft as an instrument for college education, *Journal of Health Economics* **26**(5): 877–895.

de Walque, D. (2010). Education, information, and smoking decisions: Evidence from smoking histories in the united states, 1940–2000, *Journal of Human Resources* **45**(3): 682–717.

Deaton, A. (2003). Health, inequality, and economic development, *Journal of Economic Literature* **41**: 113–158.

Der Spiegel (20.1.1964). Rauchen (cover topic), **1964**(4).

Destatis (2012). Geburten in Deutschland, *Report 2012*, Statistisches Bundesamt, Wiesbaden.

Dunifon, R., Kalil, A., Crosby, D. A. and Su, J. H. (2013). Mothers' night work and children's behavior problems, *Developmental Psychology* **49**(10): 1874–1885.

Farrell, P. and Fuchs, V. R. (1982). Schooling and health, *Journal of Health Economics* **1**(3): 217–230.

Fertig, A., Glomm, G. and Tchernis, R. (2009). The connection between maternal employment and childhood obesity: Inspecting the mechanisms, *Review of Economics of the Household* **7**(3): 227–255.

Fitzpatrick, M. D. (2010). Preschoolers enrolled and mothers at work? The effects of universal prekindergarten, *Journal of Labor Economics* **28**(1): 51–85.

Fletcher, J. M., Sindelar, J. L. and Yamaguchi, S. (2011). Cumulative effects of job characteristics on health, *Health Economics* **20**(5): 553–570.

Fuchs, V. R. (1982). Time preferences and health: An exploratory study, *in* V. R. Fuchs (ed.), *Economic Aspects of Health*, Chicago Press, pp. 93–120.

Garcia, E., Labeaga, J. M. and Ortega, C. (2006). Maternal Employment and childhood obesity in Spain, *Working Paper 2006-17*, FEDEA, Madrid.

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note, *Journal of Child Psychology and Psychiatry* **38**(5): 581–586.

Greve, J. (2011). New results on the effect of maternal work hours on children's overweight status: Does the quality of child care matter?, *Labour Economics* **18**(5): 579–590.

Grimard, F. and Parent, D. (2007). Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?, *Journal of Health Economics* **26**(5): 896–926.

Grossman, M. (2006). Chapter 10: Education and nonmarket outcomes, *in* E. Hanushek and F. Welch (ed.), *Handbook of the Economics of Education*, Elsevier, pp. 577–633.

Grossman, M. (1972). On the concept of health capital and the demand for health, *Journal of Political Economy* **80**(2): 223–255.

Grossman, M. (2015). The relationship between health and schooling: What's new?, *Nordic Journal of Health Economics* **3**(1): 7–17.

Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Ahrens, W., Eiben, G., Fernandéz-Alvira, J. M., Hadjigeorgiou, C., Henauw, T., Kovács, E., Lauria, F., Veidebaum, T., Williams, G. and Bammann, K. (2013). Maternal employment and childhood obesity – A European perspective, *Journal of Health Economics* **32**(4): 728–742.

Hartmann, J. (2006). BIBB/BauA Erwerbstätigenbefragung 2005/2006, *Feldbericht*, TNS Infratest Sozialforschung.

Heublein, U., R., J., Schmelzer, R. and Sommer, D. (2012). Die Entwicklung der Schwund- und Studienabbruchquoten an den deutschen Hochschulen. Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2010, *Forum Hochschule 3/2012*, Hochschul-Informations-System.

Himes, J. H. (2009). Challenges of accurately measuring and using BMI and other indicators of obesity in children, *Pediatrics* **124**(1): 3–22.

Hsieh, C.-S. and van Kippersluis, H. (2015). Smoking initiation: Peers and personality, *TI Discussion Paper 2015-093/V*, Tinbergen Institute.

Huisman, M., van Lenthe, F., Avendano, M. and Mackenbach, J. (2008). The contribution of job characteristics to socioeconomic inequalities in incidence of myocardial infarction, *Social Science & Medicine* **66**(11): 2240–2252.

Hunsberger, M., Formisano, A., Reisch, L. A., Bammann, K., Moreno, L., Henauw, S. d., Molnar, D., Tornaritis, M., Veidebaum, T., Siani, A. and Lissner, L. (2012). Overweight in singletons compared to children with siblings: the IDEFICS study, *Nutrition & Diabetes* **2**(7): e35.

Idler, E. L. and Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies, *Journal of Health and Social Behavior* **38**(1): 21–37.

Jensen, R. and Lleras-Muney, A. (2012). Does staying in school (and not working) prevent teen smoking and drinking?, *Journal of Health Economics* **31**(4): 644–657.

Johnson, N. J., Sorlie, P. D. and Backlund, E. (1999). The impact of specific occupation on mortality in the U.S. National Longitudinal Mortality Study, *Demography* **36**(3): 355.

Johnston, D. W., Lordan, G., Shields, M. A. and Suziedelyte, A. (2015). Education and health knowledge: Evidence from UK compulsory schooling reform, *Social Science & Medicine* **127**: 92–100.

Jürges, H., Reinhold, S. and Salm, M. (2011). Does schooling affect health behavior? Evidence from the educational expansion in Western Germany, *Economics of Education Review* **30**(5): 862–872.

Jürges, H., Thiel, L. and Börsch-Supan, A. (2016). Healthy, happy and idle: Estimating the health capacity to work at older ages in Germany, *Working Paper 21972*, National Bureau of Economic Research.

Kaikkonen, R., Rahkonen, O., Lallukka, T. and Lahelma, E. (2009). Physical and psychosocial working conditions as explanations for occupational class inequalities in self-rated health, *The European Journal of Public Health* **19**(5): 458–463.

Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign, *Administrative Science Quarterly* **24**(2): 285–308.

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

Kenkel, D. (2012). Health behaviours among young people, *in* A. M. Jones (ed.), *The Elgar Companion to Health Economics*, Edward Elgar Publishing, Cheltenham, pp. 60–69.

Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling, *Journal of Political Economy* **99**(2): 287–305.

Kenkel, D. S., Lillard, D. R. and Mathios, A. (2006). The roles of high school completion and GED receipt in smoking and obesity, *Journal of Labor Economics* **24**(3): 635–660.

Kenkel, D. S. and Sindelar, J. (2011). Economics of health behaviors and addictions: Contemporary issues and policy implications, *in* S. Glied and P. C. Smith (eds), *The Oxford Handbook of Health Economics*, Oxford University Press, New York, pp. 206–231.

Krall, E. A., Valadian, I., Dwyer, J. T. and Gardner, J. (1989). Accuracy of recalled smoking data., *American Journal of Public Health* **79**(2): 200–202.

Kroll, L. E. (2011). Konstruktion und Validierung eines allgemeinen Index für die Arbeitsbelastung in beruflichen Tätigkeiten auf Basis von ISCO-88 und KldB-92, *Methoden, Daten, Analysen* **5**(1): 63–90.

Kromeyer-Hauschild, K., Wabitsch, M., Kunze, D., Geller, F., Geiß, H. C., Hesse, V., Hippel, A. v., Jaeger, U., Johnsen, D., Korte, W., Menner, K., Müller, G., Müller, J. M., Niemann-Pilatus, A., Remer, T., Schaefer, F., Wittchen, H.-U., Zabransky, S., Zellner, K., Ziegler, A. and Hebebrand, J. (2001). Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben, *Monatsschrift Kinderheilkunde* **149**(8): 807–818.

Kurth, B.-M., Kamtsiuris, P., Hölling, H., Schlaud, M., Dölle, R., Ellert, U., Kahl, H., Knopf, H., Lange, M., Mensink, G. B., Neuhauser, H., Rosario, A. S., Scheidt-Nave, C., Schenk, L., Schlack, R., Stolzenberg, H., Thamm, M., Thierfelder, W. and Wolf, U. (2008). The challenge of comprehensively mapping children's health in a nation-wide health survey: Design of the German KiGGS-Study, *BMC Public Health* **8**: 196.

Kvalseth, T. O. (1985). Cautionary note about R^2, *The American Statistician* **39**(4): 279–285.

Lahelma, E. (2004). Pathways between socioeconomic determinants of health, *Journal of Epidemiology & Community Health* **58**(4): 327–332.

Lampert, P. D. T., Müters, S., Stolzenberg, H., Kroll, L. E., Group, K. S. et al. (2014). Messung des sozioökonomischen Status in der KiGGS-Studie, *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* **57**(7): 762–770.

Li, C., Goran, M. I., Kaur, H., Nollen, N. and Ahluwalia, J. S. (2007). Developmental trajectories of overweight during childhood: Role of early life factors, *Obesity (Silver Spring, Md.)* **15**(3): 760–771.

Li, J. and Powdthavee, N. (2015). Does more education lead to better health habits? Evidence from the school reforms in Australia, *Social Science & Medicine* **127**(0): 83–91.

Lochner, L. (2011). Nonproduction benefits of education, *in* E. A. Hanushek, F. Welch, S. Machin and L. Woessmann (eds), *Handbook of the Economics of Education*, Vol. 4, North Holland, Amsterdam, pp. 183–282.

Lundborg, P. (2013). The health returns to schooling - What can we learn from twins?, *Journal of Population Economics* **26**(2): 673–701.

Mackenbach, J. P. (2011). The English strategy to reduce health inequalities, *The Lancet* **377**(9782): 1986–1988.

Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M. and Kunt, A. E. (2008). Socioeconomic inequalities in health in 22 european countries, *New England Journal of Medicine* **358**(23): 2468–2481.

Mahler, P. (2007). I'm not fat, just too short for my weight: Family child care and obesity in Germany, *Working Paper 0707*, Socioeconomic Institute – University of Zurich.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem, *The Review of Economic Studies* **60**(3): 531–542.

Maralani, V. (2013). Educational inequalities in smoking: The role of initiation versus quitting, *Social Science & Medicine* **84**: 129–137.

Maralani, V. (2014). Understanding the links between education and smoking, *Social Science Research* **48**(0): 20–34.

McMichael, A. J., Spirtas, R. and Kupper, L. L. (1974). An epidemiologic study of mortality within a cohort of rubber workers, 1964-72, *Journal of Occupational and Environmental Medicine* **16**(7): 458–464.

Miller, D. P. (2011). Maternal work and child overweight and obesity: The importance of timing, *Journal of Family and Economic Issues* **32**(2): 204–218.

Mirowsky, J. and Ross, C. E. (2003). *Education, Social Status, and Health*, Social Institutions and Social Change, Transaction Publishers.

Mokdad, A. H., Marks, J. S., Stroup, D. F. and Gerberding, J. L. (2004). Actual causes of death in the United States, 2000, *JAMA* **291**(10): 1238–1245.

Monden, C. W. (2005). Current and lifetime exposure to working conditions. Do they explain educational differences in subjective health?, *Social Science & Medicine* **60**(11): 2465–2476.

Morefield, G. B., Ribar, D. C. and Ruhm, C. J. (2011). Occupational status and health transitions, *Working Paper Series 16794*, National Bureau of Economic Research.

Morrill, M. S. (2011). The effects of maternal employment on the health of school-age children, *Journal of Health Economics* **30**(2): 240–257.

Morrissey, T. W. (2013). Trajectories of growth in body mass index across childhood: Associations with maternal and paternal employment, *Social Science & Medicine* **95**: 60–68.

Morrissey, T. W., Dunifon, R. E. and Kalil, A. (2011). Maternal employment, work schedules, and children's body mass index, *Child Development* **82**(1): 66–81.

Möser, A., Chen, S. E., Jilcott, S. B. and Nayga, R. M. (2012). Associations between maternal employment and time spent in nutrition-related behaviours among German children and mothers, *Public Health Nutrition* **15**(7): 1256–1261.

Nie, P. and Sousa-Poza, A. (2014). Maternal employment and childhood obesity in China: Evidence from the China Health and Nutrition Survey, *Applied Economics* **46**(20): 2418–2428.

Niedhammer, I., Chastang, J.-F., David, S. and Kelleher, C. (2008). The contribution of occupational factors to social inequalities in health: Findings from the national French SUMER survey, *Social Science & Medicine* **67**(11): 1870–1881.

OECD and EU (2014). *Health at a Glance: Europe 2014*, OECD Publishing, Paris.

Pampel, F. C. (2005). Diffusion, cohort change, and social patterns of smoking, *Social Science Research* **34**(1): 117–139.

Pampel, F. C., Legleye, S., Goffette, C., Piontek, D., Kraus, L. and Khlat, M. (2015). Cohort changes in educational disparities in smoking: France, Germany and the United States, *Social Science & Medicine* **127**: 41–50.

Park, C. and Kang, C. (2008). Does education induce healthy lifestyle?, *Journal of Health Economics* **27**(6): 1516–1531.

Piontek, D., Kraus, L., Müller, S. and Pabst, A. (2010). To what extent do age, period, and cohort patterns account for time trends and social inequalities in smoking?, *SUCHT* **56**(5): 361–371.

Qiu, H., Bures, R. and Shehan, C. L. (2012). The inconsistent mediating effects of psychosocial work characteristics on the education–health relationship, *Social Science & Medicine* **75**(8): 1539–1546.

Rabe-Hesketh, S. and Skrondal, A. (2008). *Multilevel and longitudinal modeling using stata*, 2nd edn, Stata Press Publication, College Station, Texas.

Rapoport, B. and Le Bourdais, C. (2008). Parental time and working schedules, *Journal of Population Economics* **21**(4): 903–932.

Ravens-Sieberer, U. and Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytical results, *Quality of Life Research* **7**(5): 399–407.

Reinhold, S. and Jürges, H. (2012). Parental income and child health in Germany, *Health Economics* **21**(5): 562–579.

Rosenzweig, M. R. and Schultz, T. P. (1981). *Education and Household Production of Child Health*, American Statistical Association, Washington, Proceedings of the American Statistical Association, Social Statistical Section.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology* **66**(5): 688–701.

Ruhm, C. J. (2004). Parental employment and child cognitive development, *Journal of Human Resources* **XXXIX**(1): 155–192.

Ruhm, C. J. (2006). Macroeconomic conditions, health and mortality, *in* A. M. Jones (ed.), *Elgar Companion to Health Economics*, Edward Elgar Publishing, Cheltenham and UK.

Ruhm, C. J. (2008). Maternal employment and adolescent development, *Labour Economics* **15**(5): 958–983.

Schneider, J. and Beblo, M. (2010). Health at work: Indicators and determinants. A revised literature and data review for Germany, *IAB Discussion Paper 17/2010*, Nürnberg.

Schulze, A. and Mons, U. (2006). The evolution of educational inequalities in smoking: A changing relationship and a cross-over effect among German birth cohorts of 1921–70, *Addiction* **101**(7): 1051–1056.

Sekine, M., Chandola, T., Martikainen, P., Marmot, M. and Kagamimori, S. (2009). Socioeconomic inequalities in physical and mental functioning of British, Finnish, and Japanese civil servants: Role of job demand, control, and work hours, *Social Science & Medicine* **69**(10): 1417–1425.

Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions, *Journal of Occupational Health Psychology* **1**(1): 27.

Simon, J., De Boer, J., Joung, I., Bosma, H. and Mackenbach, J. (2005). How is your health in general? A qualitative study on self-assessed health, *The European Journal of Public Health* **15**(2): 200–208.

Smith, J. P. (2004). Unraveling the SES–health connection, *Population and Development Review* **30**: 108–132.

Spiess, C. and Wrohlich, K. (2008). The parental leave benefit reform in Germany: Costs and labour market outcomes of moving towards the nordic model, *Population Research and Policy Review* **27**(5): 575–591.

Terza, J. V., Basu, A. and Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling, *Journal of Health Economics* **27**(3): 531–543.

Vedøy, T. F. (2014). Tracing the cigarette epidemic: An age-period-cohort study of education, gender and smoking using a pseudo-panel approach, *Social Science Research* **48**(0): 35–47.

von Hinke Kessler Scholder, S. (2008). Maternal employment and overweight children: Does timing matter?, *Health Economics* **17**(8): 889–906.

Von Hippel, P. T. and Lynch, J. L. (2014). Why are educated adults slim - Causation or selection?, *Social Science & Medicine* **105**: 131–139.

Wagner, G., Frick, J. and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) Scope, Evolution and Enhancements, *Schmollers Jahrbuch: Zeitschrift fr Wirtschafts- und Sozialwissenschaften* **127**(1): 139–169.

Warren, J. R., Hoonakker, P., Carayon, P. and Brand, J. E. (2004). Job characteristics as mediators in SES–health relationships, *Social Science & Medicine* **59**(7): 1367–1378.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*, 2nd ed. edn, MIT Press, Cambridge and Mass.

Wrohlich, K. (2011). Labor supply and child care choices in a rationed child care market, *Discussion Papers 1169*, Deutsches Institut für Wirtschaftsforschung (DIW) Berlin.

Ziol-Guest, K. M. (2014). A commentary on “maternal work and children’s diet, activity, and obesity”, *Social Science & Medicine* **107**: 205–208.

Ziol-Guest, K. M., Dunifon, R. E. and Kalil, A. (2013). Parental employment and children’s body weight: Mothers, others, and mechanisms, *Social Science & Medicine* **95**: 52–59.