Four Essays on the Socio-Economic Causes and Consequences of Individual Health as well as Public Health Crises

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Universität Potsdam

vorgelegt von

Daniel Graeber

Datum der Disputation: 28. Juni 2021

Erstgutachter: Professor Dr. Marco Caliendo

Zweitgutachter: Professor Dr. Daniel D. Schnitzlein

Eingereicht im April 2021

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-51517
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-515175

Acknowledgements

I am highly indebted to Marco Caliendo and Daniel D. Schnitzlein for acting as first and second supervisor, respectively, and for providing valuable feedback and guidance throughout my Ph.D. Many thanks go to all my co-authors, Alexander Kritikos, Felicitas Schikora, Daniel D. Schnitzlein and Johannes Seebauer for productive and enjoyable collaborations.

The chapters of this Ph.D. thesis gained much from many discussions with various researchers. Notably, I thank Prashant Bharadwaj, Eli Berman, Katrine Løken, Andreas Tveito, Esteban García, Mett Rasmusson and Ricardo Dahis as well as the UCSD Economics Department for constructive feedback, and for making my research stay at the UCSD such a fantastic experience.

I am also incredibly grateful for the DIW Graduate Center and the SOEP at the DIW Berlin for creating such a stimulating research environment. Many thanks also to Renke Schmacker, Melanie Koch, Christoph Halbmeier, Patrick Burauel, Matteo Targa and Johannes König for many enjoyable moments. This also includes the group "Applied Panel Analysis" and associates, in particular Maximilian Longmuir and Mattis Beckmannshagen. In addition, I am especially grateful to Carsten Schröder for his continuous support.

Moreover, I thank my parents who have always supported me along this way and without whom this would not have been possible. I also want to thank all my friends. Importantly, a big thanks goes to Marike Rüder who supported me fiercely along this journey. Lastly, I thank the German Academic Exchange service for their financial support.

Contents

List of Figures

List of Tables

General Introduction

Socio-economic differences in health are ubiquitous in modern societies. Life expectancy in the United States, for instance, ranges from 72.7 years at the bottom of the income distribution to 87.3 years at the top (Chetty et al., 2016). Similarly, the difference in life expectancy between the lowest and highest income group amount to 8.6 years in Germany (Lampert et al., 2019). These differences emerge, regardless which socio-economic dimension or health outcome one considers (e.g. Bhattacharya and Lakdawalla, 2006; Galama and van Kippersluis, 2018).

Moreover, these health differences widened in the last decades. Chetty et al. (2016) show that the annual increase in life expectancy between 2001 and 2014 was 0.20 years in the highest income quartile but just 0.08 years in the lowest. Similar developments have been documented for Germany Lampert et al. (2019).[1] These figures raise important questions about how societies evaluate and address existing socio-economic differences.

In economics, and in the social sciences more broadly, distributions of desirable outcomes such as health have traditionally been analyzed through the lens of a welfarist paradigm. The most extreme form of welfarism is the utilitarian approach, which implies an additive aggregation of individual utilities as a social objective function (Ferreira and Peragine, 2016). This approach provides very little scope for procedural considerations other than efficiency, however, and results in a tendency of the utilitarian social planner to tolerate socio-economic differences in health. In some cases, from this point of view, socio-economic differences in health could even be considered desirable. Taking as an example a society that has to allocate scarce inputs and aims to maximize health, measured in total life expectancy: If individuals with higher socio-economic status (SES) yield higher returns in health for a given input, the optimal

[1] These statistics focus on men. Similar differences have been observed for women (Chetty et al., 2016; Lampert et al., 2019).

program would allocate the most resources to individuals with a higher SES The economic literature indeed supports the notion that better educated individuals are more efficient producers of health (e.g. Grossman, 1972; Galama and van Kippersluis, 2018).

The welfarist paradigm was increasingly challenged over the second half of the last century, and it has since been replaced by the Equality of Opportunities approach to the evaluation of social justice (Ferreira and Peragine, 2016). According to this paradigm, an equitable society is one in which every individual has an equal chance of achieving a desired outcome. Such a society should compensate for differences that are rooted in differential circumstances with the aim of reducing such inequalities. Clearly, there is room for debate on which characteristics should be taken into consideration, but most Western societies agree that race, immigration, gender, ethnicity, and family background constitute an important set of circumstances that determine opportunity. Health differences that emerge as a consequence of these circumstances are therefore deemed undesirable.

The aim of this thesis is to quantify and explain how individual circumstances lead to different levels of health, and how health pandemics interact with circumstances to produce differential economic outcomes. The results will allow policy makers to design policies and programs to compensate for the differences in health that emerge due to circumstances. This is an ambitious endeavor for many reasons: First, the relationship between circumstances and health is prone to confounding from (un)observable characteristics. One example of such a characteristic is ability: If it were associated with higher SES and better health, as might be expected, simple OLS estimates would result in inconsistent estimates of the effect of SES on health. Second, very few data sets are sufficient to quantify the effect of SES on individual health. Third, health is an abstract construct, similar to ability, and as such is difficult to measure (Currie and Madrian, 1999). This thesis makes significant contributions to the literature in all three of these areas. In what follows, I first provide a short summary of these contributions along with more detailed overviews of the individual chapters. I conclude by highlighting how the findings may help social planners to address the problem of compensating for differences in health that are caused by differences in circumstances.

In Chapter 1, I offer the first estimates in the literature to date on intergenerational positional mobility in permanent health in Germany. Further, I anchor permanent health in permanent income, which, to my knowledge, has not been shown in the existing literature. This is a significant contribution, linking the emerging literature on intergenerational health mobility (Halliday et al., 2021, 2020; Andersen, 2019) to the established literature on intergenerational income mobility (Solon, 1992; Haider and Solon, 2006; Chetty et al., 2014). In Chapter 2, my co-author and I estimate the effect of maternal schooling on children's mental health in adulthood for the first time in the literature, and present important evidence on mediators between maternal schooling and children's mental health. In Chapter 3, my co-author and I provide the first estimates of the effects of hate crime on refugees' mental health. Further, we add important evidence on Becker and Rubinstein's (2011) hypothesis that individuals with higher ability are less susceptible to hate crimes by showing that in the case of refugees, it is their country-specific human capital, that is, language proficiency, that matters. Finally, in Chapter 4, we present the first estimate of the differential probability of income losses due to the COVID-19 pandemic among self-employed women relative to self-employed men. As we show, this gender gap can be explained entirely by the disproportionate sorting of self-employed women into industries that are more severely affected by the COVID-19 pandemic than others. Moreover, we are the first to present evidence on the exact mechanism by which these industry-specific differences emerge.

Chapter 1: Intergenerational health mobility in Germany In this chapter, I contribute to the economic literature by estimating the intergenerational mobility in permanent health in Germany. Up to now, the economic literature on intergenerational mobility has focused primarily on intergenerational mobility in income (e.g. Solon, 1992; Chetty et al., 2014; Bratberg et al., 2017; Corak, 2019; Mazumder, 2005; Zimmerman, 1992), occupational prestige (e.g. Long and Ferrie, 2007, 2013; Modalsli, 2017) or education (e.g. Blanden, 2013; Couch and Dunn, 1997; Alesina et al., 2021). However, despite the paramount importance of health for individuals' (economic) well-being (e.g. Grossman, 1972; Galama and van Kippersluis, 2018; Dalgaard and Strulik, 2014), economists have paid relatively little attention to quantifying the intergenerational persistence in health.

My focus on permanent health stems from the observation that it is mainly permanent rather than transitory health changes that affect individuals' earnings (Blundell et al., 2016; Kemptner, 2019; Keane et al., 2018; Britton and French, 2020). This is analogous to the role of permanent income, as shown in the literature on intergenerational income mobility (Becker and Tomes, 1979; Solon, 1992, 1999; Friedman, 1957).

There are three key reasons for the scarce literature on intergenerational mobility in permanent health: First, extensive and detailed health data are required, containing socio-economic information spanning long periods of time. Second, these data must allow the linkage of parents to their children in adulthood. Third, health is a latent concept, similar to ability, which presents us with the formidable task of finding a good proxy for health.

I solve these three issues in the first chapter of this thesis by describing the intergenerational positional mobility in permanent health for Germany based on the Socio-Economic Panel (SOEP), which includes more than 25 years of rich health information on parents and their adult children. Furthermore, I summarize these multidimensional health indicators into a single index, similar to the approach taken in the literature on health and earnings (Blundell et al., 2021; Britton and French, 2020). Because this health index does not exhibit a natural scale, I anchor this metric in permanent income, which is of central interest in the economic literature on intergenerational mobility (e.g. Becker and Tomes, 1979; Solon, 1999; Adermon et al., 2019).

My analysis proceeds in three steps. In the first part of the chapter, I present estimates of intergenerational positional mobility in permanent health in Germany. Throughout, I carefully account for transitory shocks and life-cycle biases. Based on the permanent health distributions, I calculate individuals' percentile rank in the permanent health distribution and perform rank-rank regressions, a method that was pioneered by Dahl and DeLeire (2008) and popularized by Chetty et al. (2014). The resulting rank-rank slope is the central statistic describing relative positional health mobility. The estimate of the intercept provides information about children's expected health rank if their parents are located at the bottom of the health distribution.

My central findings are as follows: A 10 percentile point increase in the parents' percentile rank is associated with an expected increase of 2.32 points in the child's percentile rank. This is similar to the rank-rank slope for permanent income in Germany (Bratberg et al., 2017). However, the rank-rank slope, which is a measure of relative mobility, leads to ambiguous welfare interpretations. For this reason, I calculate measures of upward and downward mobility, which are the expected percentile ranks if the parents are located at the 25th and 75th percentile rank, respectively. The results show that upward and downward mobility in permanent health are 44.43 and 56.54, respectively.

In the second part of the chapter, I contribute to the literature by anchoring the permanent health distribution in permanent income, a common method in the literature on skill production (e.g. Cunha and Heckman, 2008; Cunha et al., 2010; Cunha, 2011; Bond and Lang, 2018). My approach provides guidance on how to overcome the lack of a natural metric in the health economics literature. This is important since studies in health economics increasingly rely on latent variables models (e.g. Andersen, 2019; Halliday and Mazumder, 2017; Halliday et al., 2020) and generic health measures such as the Short Form 12 (SF-12) questionnaire (e.g. Marcus, 2013; Eibich, 2015), the Kessler scale (e.g Adhvaryu et al., 2019), or the Center for Epidemiological Studies Depression scale (e.g. Papageorge et al., 2019; Fruehwirth et al., 2019). Every measure that has a natural metric and is correlated with health qualifies as its anchor metric (Cunha, 2011). Permanent income is one such metric (e.g. Grossman, 1972; Currie and Madrian, 1999). I show that an increase of one percentile point in the permanent health distribution is associated with an increase in permanent income of between 0.8 and 1.4% in both generations.

However, deviating from the assumption of linearity, I find evidence for strong nonlinearities in the association between permanent health and permanent income. In all generations, the association between the percentile rank in the permanent health distribution and permanent income is highly nonlinear and stronger in the bottom quintile of the distribution of permanent health. Thus, changes in permanent health are particularly consequential for individuals at the bottom of the health distribution. This points towards strong incentives to escape the bottom of the health distribution across generations: for instance, altruistic parents with higher SES who are located at the bottom of the health distribution have strong incentives to invest in their children's health. A direct implication is that having a higher SES in childhood should be associated with higher upward mobility in health. I test this hypothesis in the third part of Chapter 1, where I investigate how intergenerational health mobility interacts with parental SES. I compare children's upward and downward mobility of parents with a low SES with the health of parents with a high SES, who are located at the same percentile rank of the parental permanent health distribution, i.e., parents with the same health endowment but different socio-economic characteristics. Strikingly, I find that improvements in SES are associated with higher upward mobility in health. This is consistent with my conclusion that the high nonlinearities in the association between permanent income and health create strong incentives to escape the bottom of the health distribution. This evidence also stands in clear contrast to findings for the United States (Halliday et al., 2021), where children of parents with more "favorable" socio-economic characteristics are better off over the whole parental health distribution.

In Chapter 1, I make a number of important contributions to the literature: I am the first to estimate the intergenerational positional mobility in permanent health for Germany, and the first to anchor permanent health in permanent income, creating a bridge between the literature on health and that on income mobility. The available multidimensional measures of health in large-scale survey data often include measures of health that lack a natural scale. Anchoring such measures in a natural scale allows researchers to circumvent this issue. Finally, the focus on the rank-rank slope as a measure of intergenerational persistence in permanent health has attractive statistical features since alternative measures of persistence are confounded by changes in health inequality across generations. It is not clear a priori whether any measure of intergenerational mobility should subsume this change in inequality across generations. This is an aspect that has been neglected in the literature on intergenerational mobility in health up to now.

Chapter 2: The effect of maternal education onoffsprings' mental health In this chapter (coauthored by Daniel Schnitzlein), we estimate the effect of maternal schooling on children's mental health in adulthood. The question of the magnitude of this effect has not yet been answered in the empirical literature. This is, in our view, a major gap in the research. Mental disorders and substance use disorders accounted for 10.4% of the global burden of disease in 2010, making them the main source of years lived with disability among all disease groups (Bloom et al., 2011). In 2010, direct and indirect economic costs of mental disorders were estimated at 2.5 trillion dollars worldwide. Most worryingly, the direct and indirect costs of mental disorders are projected to double by 2030 relative to 2010 figures (Bloom et al., 2011).

Estimating the effect of maternal schooling on children's mental health in adulthood is challenging. First, the relationship of interest could be confounded by unobservable characteristics that are jointly associated with maternal schooling and children's mental health. In addition, classic measurement error in maternal years of schooling could attenuate estimates. In both cases, our OLS estimates of the relationship of interest would be inconsistent. Therefore, to identify the effect of maternal years of schooling on children's mental health in adulthood, we exploit exogenous variation in maternal years of schooling, caused by a compulsory schooling law (CSL) reform, which increased the number of compulsory years of schooling from eight to nine in the West German states. Thus, the complier group consists of mothers who attended the basic school track (*Hauptschule*).

The data we use come from the SOEP, and our main outcome is the Mental Component Summary (MCS) score, i.e., the second component of a principal component analysis of the 12 items of the Short Form-12 (SF-12) questionnaire. The MCS score is a widely used and highly predictive summary measure of individuals' mental health within the 30 days preceding the interview. An additional measure we use is the Physical Component Summary (PCS) score, the first factor of the principal component analysis of the 12 items of the SF-12 questionnaire. The PCS score is a summary measure of physical health.

Our results indicate that maternal schooling does not have any effect on children's MCS score in adulthood. This conclusion is robust to various robustness checks. Further, we are able to replicate various results from the literature, notably, the positive effect of maternal years of schooling on physical health, measured by the PCS score. Notably, further inspection of the subscales associated with the MCS and PCS score indicates that it is the self-rated health status (SRHS) and "physical functioning" that drive the results for physical health. The effect on the latter dimension has not been shown in the literature until now.

In a subsequent step, we test various mediators of the relationship between maternal years of schooling and children's mental health in adulthood.[2] We find that social

[2]It should be noted that the absence of a total effect of maternal years of schooling does not necessarily indicate the absence of any mediators of the relationship between maternal years of schooling

capital mediates the relationship between maternal years of schooling and children's mental health in adulthood. This has not been shown in the empirical literature to date. However, the implied overall effect is rather small, consistent with the reduced-form effect.

The contributions of Chapter 2 are as follows: We are the first to estimate the effect of maternal years of schooling on children's mental health in adulthood. Second, we present an additional dimension through which maternal years of schooling contribute to the overall effect on physical health. And third, we demonstrate the existence of social capital as a mediator between maternal years of schooling and children's mental health in adulthood.

Chapter 3: Hate is too great a burden to bear: Hate crimes and the mental health of refugees In this chapter (coauthored by Felicitas Schikora), we focus on migration as a further dimension of individuals' circumstances by investigating the effect of hate crime on refugees' mental health in Germany.

Quantifying the effect of hate crime on refugees' mental health is challenging. First, refugees' choice of where to take up residence in Germany is potentially endogenous. They could, for instance, try to avoid regions that are reputed to be more xenophobic. As a result, the OLS estimate of a regression of refugees' mental health on an indicator for proximity hate crime would be inconsistent. Second, such an analysis requires representative, high-quality data on refugees, their mental health, and predetermined characteristics.

We solve these problems by estimating the effect of hate crimes on refugees' mental health by means of a regression discontinuity design (RDD) in time, comparing refugees' mental health outcomes just before and after the occurrence of a hate crime in their county of residence. Our data come from the IAB-BAMF-SOEP Survey of Refugees in Germany, a representative survey that was launched to address the arrival of large numbers of refugees to the country in 2015. Our two main mental health outcomes are the MCS score and the Patient Health Questionnaire-4 (PHQ-4) score. The latter is a widely use summary score of the frequency of symptoms of depression and anxiety. The information on hate crimes, specifically attacks on refugee shelters,

and children's mental health (O'Rourke and MacKinnon, 2018).

stems from administrative information from the Federal Criminal Police Office (BKA).

We find that being close to hate crimes impairs refugees' mental health considerably. Our results indicate that the experience of a close hate crime reduces refugees' MCS score by 37% of a standard deviation. Similarly, hate crimes reduce refugees' PHQ-4 scores by 28% of a standard deviation.[3] This finding is novel to the economic literature.

Moreover, we test Becker and Rubinstein's (2011) argument that higher-educated individuals' are less affected by hate crimes. We hypothesize that the individual ability to acquire and process information plays a key role in aligning the subjective and objective probabilities of being a victim of a hate crime. We distinguish between individual levels of country-specific human capital, i.e., language proficiency, and find strong evidence that the effect sizes vary with the level of country-specific human capital: The effect sizes are larger for individuals with higher levels of country-specific human capital.

Summarizing our contributions to the literature, we are the first to estimate the effect of hate crime on refugees' mental health. Moreover, we highlight an important human capital dimension that mediates the effect of hate crime on refugees' mental health. To our knowledge, this does not exist in the economic literature to date. These are important findings, since mental health shocks have the potential to impair refugees' long-term success in their host country (Schilbach et al., 2016) and are very likely to negatively affect the next generation (Almond and Currie, 2011; Almond et al., 2018).

Chapter 4: COVID-19: A crisis of the female self-employed In this chapter (coauthored by Alexander Kritikos and Johannes Seebauer), we turn from health as an outcome to consider how the COVID-19 pandemic, a public health crisis of historic dimensions, has affected self-employed women as compared to self-employed men. As noted above, gender is an immutable circumstance according to the Inequality of Opportunity paradigm. This implies the need for social policies to eliminate socioeconomic differences that emerge based on gender.

[3]To facilitate interpretation, we inverted the scale of the PHQ-4 score. Usually, higher PHQ-4 scores indicate worse mental health. But in this application, we decided to invert the scale so that higher scores indicate better mental health. This eases interpretation together with the MCS score, for which higher values indicate better mental health.

In 2020, the COVID-19 pandemic left no domain of life untouched. In absence of any pharmaceutical measures to fight the spread of SARS-CoV-2, the virus causing the disease COVID-19, the only policies at policy makers' disposal were non-pharmaceutical interventions. These typically aimed at reducing social contact to slow the transmission rate of SARS-CoV-2. Examples of such interventions include (weak) curfews, closure of non-essential retailers, and closure of childcare facilities as well as schools (Steinmetz et al., 2020).

These measures resulted in major economic disruptions. Adams-Prassl et al. (2020) provided one of the first in-depth analyses of labor market responses to the COVID-19 pandemic. Notably, Adams-Prassl et al. (2020) report that Germany was less severely affected than the United States and the United Kingdom. In addition, they find no evidence of gender gaps in labor market shocks in the wake of the COVID-19 pandemic. Adams-Prassl et al. (2020) attribute the small labor market impact of COVID-19 to Germany's short-time work scheme, which provides companies with subsidies to compensate employees for part of their earnings loss due to pandemic-induced reductions in working hours. Importantly, they find no gender differences in job losses. One explanation for this finding is that they do not distinguish between dependent employment and self-employment.

We investigate the differential effect of the COVID-19 pandemic and associated policies on income, working time, and the likelihood of working from home for women and men in self-employment using data from the SOEP-CoV study. SOEP-CoV is a unique data source that provides us with rich information on individuals' health and labor market and family situations, among other aspects, at the onset of the COVID-19 pandemic in the second quarter of 2020. Most importantly, SOEP-CoV is a random subsample of the SOEP population. Thus, we are also able to exploit rich pre-crisis information on individuals (Kühne et al., 2020).

Our findings indicate that self-employed women are 35% more likely to have experienced income losses than their male counterparts. Our results further show that the gender gap among the self-employed is explained by the self-employed women's disproportionate representation in industries that are more severely affected by the COVID-19 pandemic. Further, we find that women are significantly more likely to have been impacted by government-imposed restrictions, e.g., the limits on store opening hours. We also find that the effect of household specialization plays a role. However, our estimates are too imprecise to confirm that this specialization is a potential driver of the observed gender gap.

Overall, my thesis provides new insights into socio-economic differences in health as well as differences in the impacts of a public health crisis that are rooted in immutable circumstances. As these differences violate the norm of equality of opportunity, social planners should aim at eradicating these differences. This thesis provides important considerations that can help guide social planners and policy makers.

The first two chapters of this thesis focus on the effects of parental background on adult health. This focus is of particular value to social planners in light of the high costs of ex-post compensation for health differences that emerge from family background. The potential expense is exemplified by the treatment of cardiovascular illnesses at older ages: Individuals with cardiovascular diseases have to change their lifestyle, take medications, and, in severe cases, leave the labor force altogether or be hospitalized. To the extent that cardiovascular diseases are rooted in conditions individuals experience during childhood, this provides scope for far less costly interventions in early stages of life.

The results presented here demonstrate that differences in individual health and capabilities are indeed rooted in early childhood conditions. Moreover, individuals' capabilities exhibit dynamic complementarities: Investments in period t depend on the level of investments in $t-1$, which makes early investments particularly efficient in the long run (Heckman, 2007; Cunha and Heckman, 2008; Cunha et al., 2010; Almond and Currie, 2011; Almond et al., 2018). And the fact that early interventions are more cost efficient than later interventions finds astonishing empirical support: Hendren and Sprung-Keyser (2020) compare 133 policy changes in the United States over the second half of the last century and conclude that the average Marginal Value of Public Funds (MVPF), defined as the individual's willingness to pay over the net government costs, is highest for direct investments in the health and education of low-income pupils. For these young people, the average MVPF is larger than 5, whereas for adults, the average MVPF ranges from 0.5 to 2. This reinforces the notion that interventions earlier in life have a higher benefit-to-costs ratio. The first two chapters contribute evidence to this public discourse.

The third chapter explores potential differences in health that are rooted in migration. Health differences between immigrants and the native population are well documented in the literature (e.g. Antecol and Bedard, 2006; Domnich et al., 2012; Jasso et al., 2004; zur Nieden and Sommer, 2016; Palloni and Arias, 2004; Razum et al., 1998; Ullmann et al., 2011; Giuntella and Mazzonna, 2015; Kennedy et al., 2006; Antman et al., 2020). The integration of refugees is of crucial importance to host societies and has been the subject of extensive research in economics (e.g. Dustmann and Glitz, 2011). There is substantial evidence that mental health differences can lead to worse economic decisions by refugees (Becker and Rubinstein, 2011; Schilbach et al., 2016) and have negative impacts on the future of their children (Almond and Currie, 2011; Almond et al., 2018). This has two important implications: First, a welcoming environment is important for refugees' future prospects. Refugees have been integrated most successfully in countries where the resident population's attitudes toward refugees were positive (Ther, 2019). Second, if hate crimes cannot be prevented entirely, it is of utmost importance that mental health services be offered to the populations who suffer from their impacts.

Chapter 4 shows how policy interventions aimed at avoiding the spread of communicable diseases can lead to differential economic outcomes between women and men. Clearly, these interventions cannot be avoided if pharmaceutical interventions such as vaccinations are lacking. However, social planners and policy makers aiming at gender equity should seek to compensate for these differences ex post. This is especially important in the case of the self-employed, who are a cornerstone of the European Economy as well as important drivers of innovation and growth (e.g., Wong et al., 2005).

CHAPTER 1

Intergenerational health mobility in Germany*

We use 25 years of rich health information to describe three features of intergenerational health mobility in Germany. First, we describe the joint permanent health distribution of the parents and their children. A ten percentile increase in parental permanent health is associated with a 2.3 percentile increase in their child's health. Second, a percentile point increase in permanent health ranks is associated with a 0.8% to 1.4% increase in permanent income for, both, children and parents, respectively. Non-linearities in the association between permanent health and income create incentives to escape the bottom of the permanent health distribution. Third, upward mobility in permanent health varies with parental socio-economic status.

*The chapter is sole authored. The corresponding paper has been revised and resubmitted to the *Journal of Human Resources*.

1.1 Introduction

The stock of health capital is an important determinant of the time an individual can allocate to welfare enhancing market and home production. In addition, a high stock of health capital exhibits consumption value (e.g. Grossman, 1972; Dalgaard and Strulik, 2014; Galama and van Kippersluis, 2018).[1] Yet, even though health is inarguably a central determinant of individual well-being and its inequality, the literature on intergenerational mobility mainly focuses on mobility in income (e.g. Solon, 1992; Chetty et al., 2014; Bratberg et al., 2017; Corak, 2019; Mazumder, 2005), occupational prestige (e.g. Long and Ferrie, 2007, 2013; Modalsli, 2017), and education (e.g. Blanden, 2013; Couch and Dunn, 1997; Alesina et al., 2021).

The reason for the scarce literature on intergenerational health mobility is threefold: First, few data sets contain rich health information in conjunction with socioeconomic information over long periods. Second, the data must allow for linking children in adulthood with their parents. Third, health is a latent concept, like ability, that is inherently difficult to measure. For instance, if we focus solely on mortality, we would discard all health conditions that are not associated with a shortened life expectancy. Moreover, if we focus on in- and outpatient care, we would discard all health conditions that do not result in medical treatment.

We solve these issues in this chapter by describing the intergenerational positional mobility in permanent health for Germany using the Socio-Economic Panel (SOEP), which provides more than 25 years of rich health information for children in adulthood and their parents. We focus on permanent health because the contemporary literature on health and earnings emphasizes that it is permanent in contrast to transitory health differences that matter (Blundell et al., 2016; Keane et al., 2018; Britton and French, 2020). Furthermore, we apply an intuitive way to capture these multidimensional data on health in a single index. Since this health index does not exhibit a natural scale, we perform an anchoring procedure to link changes in the health distribution to a metric that allows us to describe the welfare consequences of changes in health beyond the

[1] Dalgaard and Strulik (2014) deviate from the classical health-capital theory by modeling the development of health as the accumulation of health deficits over time. A further deviation of Dalgaard and Strulik (2014) is that the deficit index does not enter utility directly. The only way through which Dalgaard and Strulik (2014) hypothesize that health affects life-time utility is through expanding the individual's life expectancy and, thus, the time horizon over which individuals can consume goods.

direct consumption value of health. This anchoring metric is permanent income, which is of central interest in the economic literature on intergenerational mobility (Becker and Tomes, 1979; Solon, 1999). Thus, we also bridge the gap between the literature on economic and health mobility. We organize our analysis into three parts.

In the first part, we present estimates of intergenerational positional mobility in health for Germany. Our main analysis focuses on children born in 1945 or later as well as their parents, who are between 30 and 65 years of age. This, along with our preferred measure of mobility, helps us to account for life-cycle biases. We show that the variation in the health items of the SOEP can be explained by a single factor, i.e., latent health capital. Based on these data, we construct a continuous index reflecting the latent health capital of the respondents based on a wide range of health measurements using methods from Item Response Theory (IRT). To capture permanent health, we calculate individual level averages of the latent health status to eliminate transitory health shocks. We then run rank-rank regressions to estimate intergenerational positional mobility in permanent health. This method is pioneered by Dahl and DeLeire (2008) and stems from the literature on income mobility. We find support for this linear specification by running local linear regressions. The resulting rank-rank slope is the central statistic describing relative positional mobility in permanent health. The estimate of the intercept informs about the expected rank of the children if their parents are located at the bottom of the distribution of permanent health.

Our central findings are as follows: A 10 percentile point increase for the parents is associated with an expected increase in the child's percentile rank of 2.32 points. This is similar to the rank-rank slope for permanent income for Germany (Bratberg et al., 2017). However, the rank-rank slope, which is a measure of relative mobility, implies ambiguous welfare interpretations. For example, an increase in relative positional mobility in health may be driven by worse health for the children of healthier parents rather than improved health for the children of parents who are less healthy. The latter case would not correspond to a Pareto improvement. Therefore, based on the estimates for the intercept and the rank-rank slope, we calculate the children's expected percentile rank in the children's distribution of permanent health if the parents are located at the 25th and 75th percentile rank. We refer to these measures as up- and downward mobility. Our estimates reveal that up- and downward mobility are

44.43 and 56.54.

In the second part, we contribute to the literature by anchoring the distribution of permanent health in permanent income. This allows us to overcome the lack of a natural metric for permanent health. This method is common in the literature of skill-production (e.g. Cunha and Heckman, 2008; Cunha et al., 2010; Cunha, 2011; Bond and Lang, 2018). Thus, we provide guidance on how to overcome the lack of a natural metric for the health economics literature. This is important since studies in health economics increasingly rely on latent variables models (e.g. Andersen, 2019; Halliday and Mazumder, 2017; Halliday et al., 2020) and generic health measurements such as the Short-Form 12 questionnaire (e.g. Marcus, 2013; Eibich, 2015), the Kessler Scale (e.g Adhvaryu et al., 2019), or the Center for Epidemiological Studies Depression Scale (e.g. Papageorge et al., 2019; Fruehwirth et al., 2019). Every measure that has (1.) a natural metric and (2.) is correlated with health qualifies as anchor metric (Cunha, 2011). Permanent income is such a metric (e.g. Grossman, 1972; Currie and Madrian, 1999). We show that an increase of a percentile point in the distribution of permanent health is associated with a 1.3% and 0.8% increase in permanent income for daughters and sons, respectively. For parents, we estimate that these associations correspond to 0.8% and 1.4% for mothers and fathers, respectively.[2]

However, deviating from the assumption of linearity, we find evidence for strong non-linearities in the association between permanent health and permanent income. In all generations, the association between the percentile rank in the distribution of permanent health and permanent income is highly non-linear and stronger in the bottom quintile of the distribution of permanent health. Thus, changes in permanent health are particularly consequential for individuals at the bottom of the health distribution. This points toward strong incentives to escape the bottom of the health distribution across generations. Therefore, altruistic parents with higher socioeconomic status (SES), who are located at the bottom of the health distribution, have strong incentives to invest in their children's health. A direct implication is that a more advantageous socioeconomic background of children should be associated with higher upward mobility in health. We test this hypothesis in the third part.

[2]The mean of the permanent income, measured in 2010 Euros, is 18,301.58 and 30,342.46 for daughters and sons, respectively. For parents, the mean of the permanent income is 15,547.78 for mothers and 35,478.19 for fathers.

Thus, in the third part, we investigate how intergenerational health mobility interacts with the parental socioeconomic background. For this, we compare children's up- and downward mobility with respect to the health of their parents, who are located at the same percentile rank of the parental distribution of permanent health, i.e. parents with the same health endowment, but have different socioeconomic characteristics. Strikingly, we find that improvements in the socioeconomic background are associated with higher upward mobility in health.[3] This is consistent with our conjecture that the high non-linearities in the association between permanent income and health creates strong incentives to escape the bottom of the health distribution. The evidence also stands in clear contrast to findings for the U.S. (Halliday et al., 2021), where children of parents with more "favorable" socioeconomic characteristics are better off across the entire parental health distribution.[4]

Our study relates primarily to the burgeoning literature on intergenerational mobility in health: Halliday et al. (2021), Halliday et al. (2020), and Fletcher and Jajtner (2019) estimate the intergenerational positional mobility in health in the U.S. using the Panel Study of Income Dynamics and the National Longitudinal Study of Adolescent to Adult Health. All three of these studies focus on self-rated health. Halliday et al. (2021) estimate a rank persistence in health of about 0.261 for the full sample. Halliday et al. (2020) build on Halliday et al. (2021) and apply a non-linear latent variable model using the self-rated health status of the individuals. They estimate a rank persistence across generations of about 0.281.[5] Fletcher and Jajtner (2019) estimate a rank persistence of about 0.174.[6] Our work differs from that of Halliday et al. (2021), Halliday et al. (2020), and Fletcher and Jajtner (2019) by considering a wider range of health outcomes. In addition, we also employ a non-linear latent variable framework, like Halliday et al. (2020), but use more health outcomes, which allows us to provide a richer characterization of the health distribution.

[3]In six out of eight cases, or 75%, differences in health mobility are characterized by higher upward mobility in health.

[4]An important caveat of this comparison are differences in methodology and cohorts under consideration.

[5]For rank-rank slopes, Halliday et al. (2021) find no differences in the estimand using either self-rated health or the non-linear latent variable model based on self-rated health. The reason is that the latter is only a positive monotone transformation of self-rated health and rank correlations are invariant to monotone transformations (Halliday et al., 2021).

[6]Fletcher and Jajtner (2019) emphasize that their estimates are heavily attenuated since parent's health is only observed once.

Further evidence stems from Andersen (2019), who estimates the intergenerational mobility of health in Denmark using administrative data on hospitalizations and general practitioner visits. Andersen (2019) characterizes the health distribution by the first principal component derived from this health information. Based on that metric, Andersen (2019) estimates rank correlations ranging from 0.112 to 0.145. In contrast to Andersen (2019), we also use subjective assessments of health, thus allowing us to consider differences in health that do not result in immediate treatment but affect individual welfare.

In addition, Germany constitutes an interesting country case in this literature. The U.S. is a country characterized by a mixture of public and private health care providers as well as high income inequality and immobility. At the opposite end, Denmark is typically described as a country with universal public health care as well as low income inequality and immobility. Along these dimensions, Germany is located in between the U.S. and Denmark (Corak, 2013). Our estimates show that Germany also ranks between the U.S. and Denmark when it comes to health mobility, constituting a new stylized fact. However, since all three of these studies rely on different health measures and cohorts, this comparison must be made with caution.

Moreover, our study also relates to the economic literature on intergenerational associations in health outcomes such as birth weight (Currie and Moretti, 2007), mental health (Johnston et al., 2013), longevity (Ahlburg, 1998; Björkegren et al., 2019; Hong and Park, 2015; Lach et al., 2006), asthma (Thompson, 2017), and self-rated health (Kim et al., 2015; Pascual and Cantarero, 2009). Compared to all these studies, we consider a broad measure of health capital instead of a single specific expression of it.[7]

For Germany, Coneus and Spiess (2012) estimate the intergenerational health as-

[7] A common theme in the literature on the intergenerational transmission of health is to what extent health is genetically determined by parents. The accumulated evidence so far is highly ambiguous. Thompson (2017) concludes that pre-birth factors account for 20-30% of the intergenerational associations in chronic conditions. However, adoption studies, such as Thompson (2017), can only plausibly distinguish between pre- and post-birth factors. For instance, a large literature on the effects of in-utero exposure to adverse conditions shows that long-run health is malleable during the fetal period (Almond et al., 2018). Fletcher and Jajtner (2019) conclude that their health mobility estimates are attenuated by 32% for self-rated health in the adoptee sample. On the opposite end, using adoptee samples, Classen and Thompson (2016) and Björkegren et al. (2019) provide evidence that BMI and mortality are largely determined by pre-birth factors, respectively. In contrast, in the genetic literature, estimates of the genetic heritability of longevity range from 15 to 30% and evidence emphasizes that this figure is inflated by a factor of up to three by positive assortative mating of the parents (Ruby et al., 2018). In conclusion, we argue that approximately 70% percent of the children's health is determined by the family environment. This provides a large scope for policy interventions.

sociation between children up to age two and their parents. Our study adds to the evidence on this important topic from Coneus and Spiess (2012) in three important ways: First, we concentrate on children in adulthood. Second, in contrast to Coneus and Spiess (2012), our health measures are reported by the children themselves.[8] Third, and this point applies to all studies on intergenerational health associations, our measure of health mobility avoids ambiguous welfare implications in two ways: First, from simple intergenerational health associations, we are not able to conclude whether changes over time or across groups correspond to Pareto improvements or not. The second aspect is related to the interpretation of intergenerational health associations. Standard OLS regressions of children's health on their parental health outcome varies with varying degrees of health inequality across generations. This is not the case for rank-rank slopes, which are scale invariant.

Further, this study relates also to the more established literature on intergenerational income mobility. Existing studies in economics focus on relative income mobility, estimating intergenerational earnings elasticities (e.g. Solon, 1992; Mazumder, 2005; Haider and Solon, 2006; Schnitzlein, 2016). A second generation of the literature focuses on positional mobility in income (e.g. Dahl and DeLeire, 2008; Chetty et al., 2014; Bratberg et al., 2017; Corak, 2019; Markussen and Røed, 2020; Bell et al., 2018; Blundell and Risa, 2019). We relate to this literature by estimating rank-rank regressions.

1.2 Methodology

1.2.1 Measuring permanent health

We focus on *permanent health* because of its relevance for individuals' earnings and life-time utility. Similarly, the focus on permanent income in the literature on intergenerational income mobility is motivated by the importance of permanent income for individuals' consumption decisions and, thus, life-time utility (Becker and Tomes, 1979; Solon, 1992, 1999; Friedman, 1957).

Existing life-cycle models that analyze the role of health on the evolution of earnings

[8] Parental reports of child health could bias estimates of the intergenerational health associations by either systematic reporting differences between high and low SES individuals or the fact that low SES is potentially associated with undiagnosed health conditions of the children (Case et al., 2002).

or labor supply over the life-cycle prominently distinguish between permanent and transitory health shocks (Blundell et al., 2016; Kemptner, 2019; Keane et al., 2018; Britton and French, 2020). These studies show that permanent health changes matter significantly more than transitory health changes. For instance, individuals would not expect large earnings penalties in response to a broken arm, i.e., a transitory shock to health. But permanent health shocks, like cardiovascular diseases, which are potentially predictable, since obesity or low physical activity predispose individuals to an elevated risk of cardiovascular diseases, are associated with large negative effects on earnings (e.g. Keane et al., 2018; Blundell et al., 2016).

Clearly, in these models, one could easily think of health as a dimension of human capital. However, health affects individuals' welfare in a multitude of ways (Grossman, 1972), many of which are different from those of education or experience, two typical proxies of human capital. One way, which is quite similar to education, is that health is associated with higher productivity. However, the level of education could also depend on permanent health. If the time-horizon over which individuals accrue returns to education is shortened by worse health prospects, the propensity to invest into human capital may decrease (Ben-Porath, 1967). Moreover, if individuals are on sick leave, they acquire less work experience and, thus, their future earnings might decrease as a consequence (Keane et al., 2018). In addition, health also has a direct consumption value (Grossman, 1972; Britton and French, 2020; Galama and van Kippersluis, 2018). Lastly, one could easily imagine a situation in which health not only enters utility directly, but in which the consumption value of other goods interact with the individuals' health status.

In the literature on health and life-cycle labor supply or earnings, authors typically summarize health in a single index, either by relying on the self-reported health status or summarizing the available health information in a summary index, e.g. a principal component analysis of a group of health items (French, 2005; French and Jones, 2011; De Nardi et al., 2017; Braun et al., 2017; Blundell et al., 2021). Since health is often measured with (classical) measurement errors, more health proxies typically lead to improved estimates of earnings or employment elasticities with respect to health (Blau and Gilleskie, 2001; Blundell et al., 2021; Britton and French, 2020). Importantly, Blundell et al. (2021) find that a single index can indeed capture important health variations for employment.

Consequently, we first summarize health in a single index, relying on item response theory (IRT). While the intuition is similar to a factor analysis, i.e., a common trait explains the common variation across items, we believe that IRT improves upon commonly applied factor analyses since it explicitly accommodates the discrete and finite nature of our data, i.e., the non-linear association between the trait and the items. Other applications of IRT in economics include Ronda (2016) and Del Boca et al. (2019). To be more specific, we use the Graded Response Model (GRM) suggested by Samejima (1969), which is appropriate for multidimensional ordinal items. Details on the method are depicted in Section 1.A of the appendix.

However, contemporaneous observations of health are only an imperfect measure for permanent health. If we do not account properly for transitory health shocks, we would expect that any coefficient of a linear regression of children's on parents' contemporaneous health status suffers from attenuation bias (Hausman, 2001; Solon, 1992). In addition, we have to account for biases that could arise due to potential heterogeneous changes of health over the life-cycle (Galama and van Kippersluis, 2018; Haider and Solon, 2006). To accommodate for the presence of transitory shocks, we take the average of individuals' health observations. In Section 1.2.2, we explain how we address potential life-cycle biases.

1.2.2 Rank mobility measures

Formally, rank mobility measures are estimated as the intercept and slope of the following linear projection:

$$r_{1iz} = \delta + \zeta r_{0z} + \eta_{zi}. \tag{1.1}$$

In Equation 1.1, r_{1iz} and r_{0z} are the percentile rank in the distribution of permanent health of child i and parents in family z, respectively. By construction, the rank of the parents is exogenous, that is $E[r_{0z} \mid \eta_{zi}] = 0$. This rules out unobserved factors that jointly determine the parents' and child's health rank and that would bias our estimates.

Then, the estimate of the intercept δ is the expected percentile rank of a child in the children's distribution of permanent health whose parents are at the bottom of the parental distribution of permanent health. The rank-rank slope ζ reflects the relative positional persistence in permanent health across generations. That is, the rank-rank slope, multiplied by 100, indicates the expected difference in the children's percentile ranks of parents who are located at the bottom and the top of the parental distribution of permanent health. Therefore, the scalar $1 - \zeta$ reflects the degree of relative positional mobility in health.

Rank-rank regressions are very popular in the literature on economic mobility (e.g. Dahl and DeLeire, 2008; Chetty et al., 2014; Bratberg et al., 2017; Corak, 2019; Bell et al., 2018; Blundell and Risa, 2019). Four reasons underlie the popularity of rank-rank regressions: First, positional mobility measures are well suited for welfare comparisons. For instance, if intergenerational health associations change over time, it is not clear whether this change corresponds to a Pareto improvement or not. As an example, suppose that the intergenerational health association decreases over time or across groups. In this case, we do not know whether the narrowing of this health gap occurs due to the children from the family with the worse health status improving or because the health status of the children of the family with the better health status deteriorates across generations. The latter case would not correspond to a Pareto improvement. Clearly, similar considerations apply to rank-rank slopes, which are also measures of relative mobility. But the estimates of the intercept and the slope of Equation 1.1 allows us to circumvent the problem of ambiguous welfare implications by calculating measures for absolute intergenerational rank mobility in health, similar to e.g. Chetty et al. (2014) or Halliday et al. (2021). Thus, we calculate the expected percentile rank in the distribution of permanent health of a child stemming from a family whose percentile rank in the distribution of permanent health is $r_{0z} \in \{25, 75\}$. We refer to these measures as absolute up- and downward mobility, respectively.

Second, every intergenerational health association depends highly on the cross-sectional inequality in the health outcome in the children's and parents' generation. To see this, the OLS coefficient of a bivariate regression of Y_{1iz} on Y_{0z} can be decomposed as follows:

$$b_{ols} = \frac{Cov(Y_{1iz}, Y_{0z})}{Var(Y_{0z})} = \frac{Cov(Y_{1iz}, Y_{0z})}{\sigma_0 \sigma_1} \frac{\sigma_1}{\sigma_0} = Corr(Y_{1iz}, Y_{0z}) \frac{\sigma_1}{\sigma_0}, \tag{1.2}$$

with σ_0 and σ_1 being the standard deviation in the health outcome in the parents' and children's generation. Further, $Cov(\cdot)$ and $Corr(\cdot)$ correspond to the covariance and correlation. For a fixed correlation in health outcomes across generations, a doubling of the cross-sectional inequality across generations doubles the intergenerational health association. This change would also increase differences in health outcomes between individuals in the children's generation and the associated consumption possibilities. Without normative foundations, it is not clear whether a measure of intergenerational health association should capture this or not.[9] In contrast, rank mobility measures are invariant in the scale of the underlying outcome.

The third reason for the popularity of rank-rank regressions is the fact that the estimates have more desirable statistical properties than intergenerational health associations. For instance, the variance of the true percentile rank and estimated percentile rank in the respective distributions of permanent health are equal by definition. Consequently, attenuation bias due to i.i.d. shocks is less of a concern (Nybom and Stuhler, 2017). Nevertheless, we show that i.i.d. health shocks could bias our estimates and that taking individual time averages is a remedy to this.

Fourth, starting from age 30, estimates of rank-rank slopes tend to show no life-cycle biases in the case of permanent income in Sweden (Nybom and Stuhler, 2017). In Section 1.B of the appendix, we depict the life-cycle properties of the latent health status for high and low SES individuals. Like earnings, early health observations could lead to biases. However, after the age of 30, a clear ordering emerges. Therefore, we restrict our sample to the age 30-65. Moreover, the inequality in health increases with age. This could also cause life-cycle biases. But rank-rank slopes are invariant to mean preserving spreads. This would not be the case for OLS estimates. Further, in Section 1.5.3, we test the robustness of our estimates to life-cycle biases and can reject the presence of life-cycle biases.

We calculate the percentile ranks in permanent health separately for all genders and generations. Before that, we partial out a second order polynomial of age as well as year of birth fixed effects for males and females as well as the child and parent generations. Lastly, we calculate the percentile ranks separately. In addition, we average the latent health status of both parents and partial out quadratic age terms

[9] Landersø and Heckman (2017) put this argument forward for the case of intergenerational income mobility.

and year of birth fixed effects for both parents as well as indicators that indicate whether the mother or the father is missing and the share of observations contributed by the mother. If the father or mother is missing, we set the respective age and year of birth equal to zero. Then, we calculate the respective permanent health of the parents jointly based on this latent health status.[10]

1.2.3 Anchoring health in a natural metric

After measuring the degree of health mobility, we still do not know how to interpret changes in the distribution of permanent health. Therefore, we anchor permanent health in an anchoring metric that exhibits a natural scale. Every outcome that (1.) exhibits a natural metric and (2.) is correlated with permanent health qualifies as anchoring metric (Cunha, 2011). We use permanent income as such a natural metric. Permanent income is (1.) measured in Euro and health is (2.) correlated with earnings.

In our case, the anchoring equation takes the form

$$y_i = \alpha + \gamma r_i + \phi_i. \tag{1.3}$$

In Equation 1.3, y_i is permanent income, adjusted for age and year of birth, and γr_i is referred to as the anchoring function (Cunha, 2011). However, the anchoring Equation 1.3 implies a linear relationship between the percentile rank in the distribution of permanent health and permanent income. As we show in Section 1.4, the relationship between the percentile rank in the distribution of permanent health and permanent income is non-linear. Thus, we also display a nonparametric specification.[11]

[10] We emphasize that health is not equivalent to earnings if we focus on the permanent health of the parents jointly. Consequently, the interpretation might change accordingly. However, many processes adhere to a regression to the mean. As an example, children tend to be of average parental height (Tanner et al., 1970). Therefore, we believe that the average health is indeed of relevance for children's health in our setting. However, where necessary, we always show also the estimates separately for all combinations of children and parents.

[11] We only show OLS associations of the underlying relationship. However, we argue that the estimated association corresponds to an upper bound. The reasons for that are twofold: First, the presence of a justification bias could bias our estimates downward (e.g. Blundell et al., 2021; Currie and Madrian, 1999). The explanation is that individuals who work fewer hours or do not work at all could be hypothesized as justifying this reduction of labor supply by their poor health status. If this is the case, we would expect that any association between subjective health proxies and labor supply or earnings is biased upwards. Second, the existence of classical measurement error in health measurements could attenuate OLS estimates, biasing estimates of the underlying relationship downwards. We discuss the relevance of classical measurement error in Section 1.5 and conclude that

1.3 Data

We use 25 waves of the SOEP to estimate the intergenerational mobility in permanent health. The SOEP is a representative panel of households in Germany that is administered to individuals and households annually since 1984. The SOEP contains rich information on occupational biographies, education, household composition, and health, among others. Today, about 15,000 households and 30,000 persons participate in the SOEP survey.[12] For more detailed information, see Göbel et al. (2018).

Most important for our study, we can link parents and their adult children in the SOEP. Children in each household of the SOEP are surveyed first when they turn eleven or twelve years old and followed thereafter, even if they leave the parental household and form new households.[13] Thus, we are able to link parents with their adult children, even if the children no longer live in the parental household.

For the IRT model to summarize health, we make use of all health items that are administered from 1992 through 2017 in a consistent way.[14] There exists no comparable data which contains consistent health information over so many years in Germany. These items are:

- The self-rated health status,
- satisfaction with health,
- number of doctor visits within the last three months,
- number of hospital admissions in the previous year,[15] and

measurement error is present and that individual time averages account well for classical measurement error. Thus, we face two sources of bias that work in opposite directions. However, since we account for classical measurement, we are left with justification bias as the only source of bias. Therefore, we argue that our estimates are either not biased or biased upward. Therefore, we conclude that our estimate represents an upper bound.

[12]We use SOEPv34. DOI: 10.5684/soep.v34.

[13]Until 2013, children in each household were surveyed first in the year in which they turned 17 years old. Since 2014, the SOEP also administers questionnaires to individuals aged eleven and twelve.

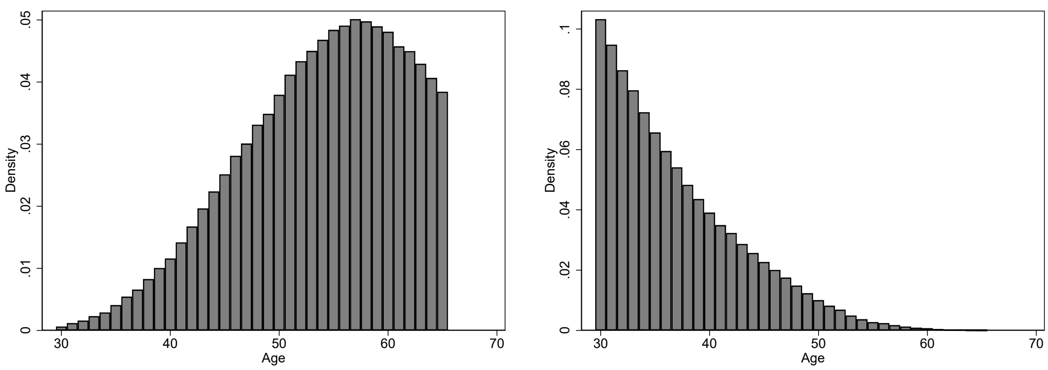
[14]We do not use the 1993 wave since the self-rated health status was not inferred in 1993.

[15]We use hospital visits of the previous year since most interviews are conducted in the first half of the year. Therefore, we argue that hospital visits of the previous year are more reflective of the health status at the time of the interview then the number of hospital visits in the contemporaneous survey year.

- the degree of disability or reduced earnings capacity[16] as assessed by a doctor.[17]

Detailed information on the health items and their operationalization as well as the IRT analysis are available in Table 1.D.1 and Section 1.C in the appendix. The age distribution of the final sample is displayed in Figure 1.1. Figure 1.2 displays the unadjusted and adjusted distribution of permanent health of the children and their parents, respectively. The unadjusted distributions of permanent health in Figure 1.2a suggest that the children have better permanent health, on average, than their parents. However, the difference is accounted for completely by age and year of birth fixed effects, as the adjusted distributions permanent health in Figure 1.2b suggest. Based on this permanent health measures, we calculate the respective percentile rank in the distributions of permanent health. The summary statistics for our main sample are displayed in Table 1.1.

Figure 1.1: Age distributions



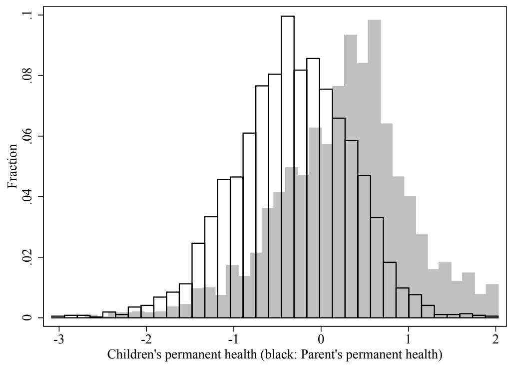
(a) Parents (b) Children

Note: Figures 1.1a and 1.1b display the age distribution in the parent and child sample.

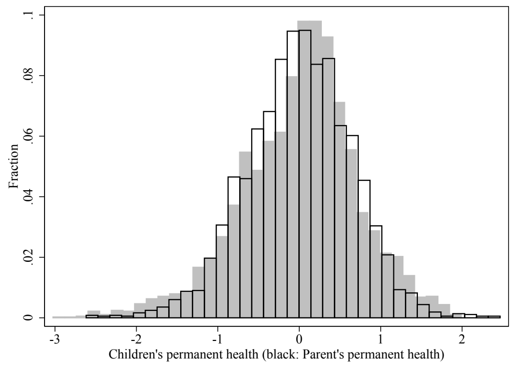
[16] In the U.S., individuals who apply for benefits from the Social Security Disability Insurance program, must be unemployed or have very low earnings. In fact these earnings have to be lower than the substantial gainful activity threshold, for at least 5 months before the receipt of the benefits can occur (e.g. Hosseini et al., 2021; Social Security Disability Insurance, 2021). This reinforces any negative correlation between application status and earnings. In contrast, there exists no formal earnings threshold in Germany. On the contrary, applicants have to have contributed to the German statutory pension insurance scheme for at least three years in the five years prior to the application for retirement benefits because of reduced earnings capacity (Deutsche Rentenversicherung, 2021). Therefore, there exists no formal earnings threshold. We imputed zeros for individuals who reported the absence of disabilities or reduced earnings capacities.

[17] Table 1.D.1 includes detailed information on the health items and on our recoding.

Figure 1.2: Distribution of permanent health



(a) Unadjusted



(b) Adjusted

Note: Figures 1.2a and 1.2b display the unadjusted and adjusted distribution of permanent health of parents and their children, respectively. Higher value correspond to better permanent health. Bars without a filling color correspond to the children's permanent health. Grey bars correspond the parent's permanent health. Figure 1.2a displays the unadjusted distributions of permanent health. Figure 1.2a displays the distributions of permanent health, adjusted for a second order polynomial in age and year of birth fixed effects.

Table 1.1: Summary statistics

	Parents		Children	
	Father (1)	Mother (2)	Son (3)	Daughter (4)
Outcomes:				
Permanent health (standard deviations)	-0.308	-0.331	0.279	0.218
	(0.766)	(0.752)	(0.717)	(0.751)
Years of health measurement	11.742	12.884	8.698	8.365
	(6.784)	(7.091)	(6.800)	(6.231)
Permanent income (2010 Euros)	30235.980	11595.393	32096.969	18668.463
	(28419.007)	(14202.182)	(20552.206)	(15162.707)
Health items:				
Self-rated health status	2.862	2.900	2.312	2.363
	(0.728)	(0.715)	(0.653)	(0.689)
Satisfaction with health	4.929	4.935	3.847	3.963
	(1.757)	(1.716)	(1.570)	(1.634)
Degree of disability	10.999	8.172	3.559	3.013
	(21.194)	(18.719)	(14.548)	(13.565)
More than 3 doctor visits last 3 months	0.248	0.282	0.115	0.215
At least 2 hospital visits in previous year	0.034	0.029	0.015	0.024
Additional characteristics				
Age	55.937	54.690	34.478	34.189
	(5.835)	(6.015)	(3.927)	(3.664)
Year of birth	1944.758	1946.860	1972.115	1974.077
	(8.603)	(8.985)	(7.959)	(7.408)
Number of individuals	3090	3536	2012	1643

Note: Table 1.1 displays summary statistics of the sample for the main analysis. Standard deviations are in parantheses.

We also construct a subsample to investigate the influence of the parental socioeconomic characteristics on the degree of intergenerational health mobility.[18] For this, we restrict the parental observations to provide information on the three most important proxies for the SES (Krieger et al., 1997) and the migration background of the parents. The three proxies for the parents' SES are parental education, the individual time average of the occupational prestige score, and permanent income.

The educational background is captured by the school leaving degrees.[19] For our analysis, we collapse the school leaving degree into two categories: The first category consists of individuals with no or a basic school leaving degree. The second category consists of individuals with an intermediate or high school leaving degree.

Permanent income is calculated as the individual average of the yearly labor earnings. The yearly labor earnings comprises wages and salary from all employment and self-employment as well as income from bonuses, overtime, and profit-sharing.[20] We partial out a second order polynomial of age and year of birth fixed effects from the logarithm of yearly labor earnings and calculate the individual time average of the parents between the age 30 and 65.[21] Our analysis then compares individuals whose parents have permanent income above and below the median.[22]

Occupational prestige is summarized by the Magnitude-Prestige Scale (MPS), developed by Wegener (1988). The advantage of the MPS over other prestige scales is that the MPS based on subjective assessments of the social prestige of occupations in

[18]We rely on a separate sample to avoid selection bias since it can be hypothesized that individuals with non-missing information on background characteristics are different than the whole sample. This is already reflected in the higher health status, on average, and higher labor earnings for the parents in our subsample.

[19]In Germany, for the generations under consideration, tracking typically starts after grade four. Children are then allocated to three different school tracks, according to their ability, as reflected in the children's GPA. Children with the lowest school grades are allocated to the basic school ("Hauptschule"), preparing the students for vocational education. Students with intermediate grades are allocated to the intermediate school ("Realschule"), comprising a more academic curriculum then the basic school, preparing their students for more demanding vocational training. The best students are typically allocated to the high school ("Gymnasium"), preparing the students for an academic education.

[20]Monthly labor earnings stem from the Cross National Equivalence Files, an international project that provides internationally harmonized household panels. For further details, see Frick et al. (2007).

[21]We proceed this way separately for each gender and generation. To calculate the joint permanent income of the parents, we proceed similarly to permanent health.

[22]Haider and Solon (2006) and Nybom and Stuhler (2016) highlight the relevance of life-cycle biases in the approximation of permanent income. The parents in our analysis are not in the recommended age range of 30 to 45 for the approximation of life-time earnings. However, since the median is an order statistic, we are confident that a median split of this measure avoids any life-cycle related problems with this measure since the median is robust to any spread of the distribution.

Germany. Thus, the MPS captures information beyond education and income. Again, we calculate the individual time average of the MPS of parents of age 30 through the age of 65.

Lastly, the parents' migration background is summarized by an indicator that is equal to one if the respective parent has a direct migration background, e.g., if a parent is born outside of Germany. We show the corresponding summary statistics of the sample for the heterogeneity analysis in Table 1.2.

Table 1.2: Summary statistics for the sample for the analysis of the interaction of the parents' socioeconomic background with health mobility

	Parents		Children	
	Father (1)	Mother (2)	Son (3)	Daughter (4)
Outcomes:				
Permanent health (standard deviations)	-0.275	-0.273	0.294	0.227
	(0.732)	(0.726)	(0.705)	(0.738)
Years of health measurement	12.251	13.748	8.614	8.249
	(6.677)	(6.914)	(6.757)	(6.134)
Health items:				
Self-rated health status	2.829	2.841	2.295	2.353
	(0.696)	(0.692)	(0.642)	(0.678)
Satisfaction with health	4.866	4.837	3.824	3.943
	(1.692)	(1.639)	(1.555)	(1.605)
Degree of disability	9.583	6.866	3.196	2.709
	(18.983)	(15.972)	(13.502)	(12.608)
More than 3 doctor visits last 3 months	0.242	0.271	0.113	0.214
At least 2 hospital visits in previous year	0.028	0.026	0.014	0.024
Additional characteristics:				
Age	55.686	53.900	34.275	34.004
	(5.677)	(5.815)	(3.712)	(3.433)
Year of birth	1945.042	1948.417	1972.579	1974.419
	(8.497)	(8.538)	(7.819)	(7.177)
Parental background characteristics:				
Basic school leaving degree or less	0.430	0.423		
Intermediate school leaving degree	0.364	0.448		
Academic school leaving degree	0.206	0.129		
Occupational prestige	58.824	56.319		
	(29.140)	(23.808)		
Migration Background	0.226	0.192		
Permanent income (2010 Euros)	32386.796	15126.383		
	(28231.922)	(14480.915)		
Number of individuals	2853	2703	1822	1545

Note: Table 1.2 displays summary statistics of the sample for the analysis of the interaction of the parents' socioeconomic background with health mobility. Standard deviations are in parantheses.

1.4 Main results

In Table 1.3, we display the results of the rank-rank regressions and the up- and downward mobility in permanent health. Throughout, robust standard errors are clustered on the family level. In addition, Figure 1.3 displays the rank-rank regression for children and their parents jointly. The rank-rank slope of the sample combining both parents and all children is 0.232. That is, if two children's parents are 10 percentile points apart in the parental permanent health distribution, this gap is expected to decrease 2.32 percentile ranks in the children's permanent health distribution. From Figure 1.3, it becomes immediately apparent that the relationship between children's and parent's percentile rank is indeed linear. In Section 1.5, we provide further evidence for the linearity assumption. Further, the rank-rank slope is 0.219 for the mother-son and 0.233 for the mother-daughter relation. Lastly, the rank-rank slopes are 0.193 and 0.198 for the father-son and father-daughter relation, respectively.

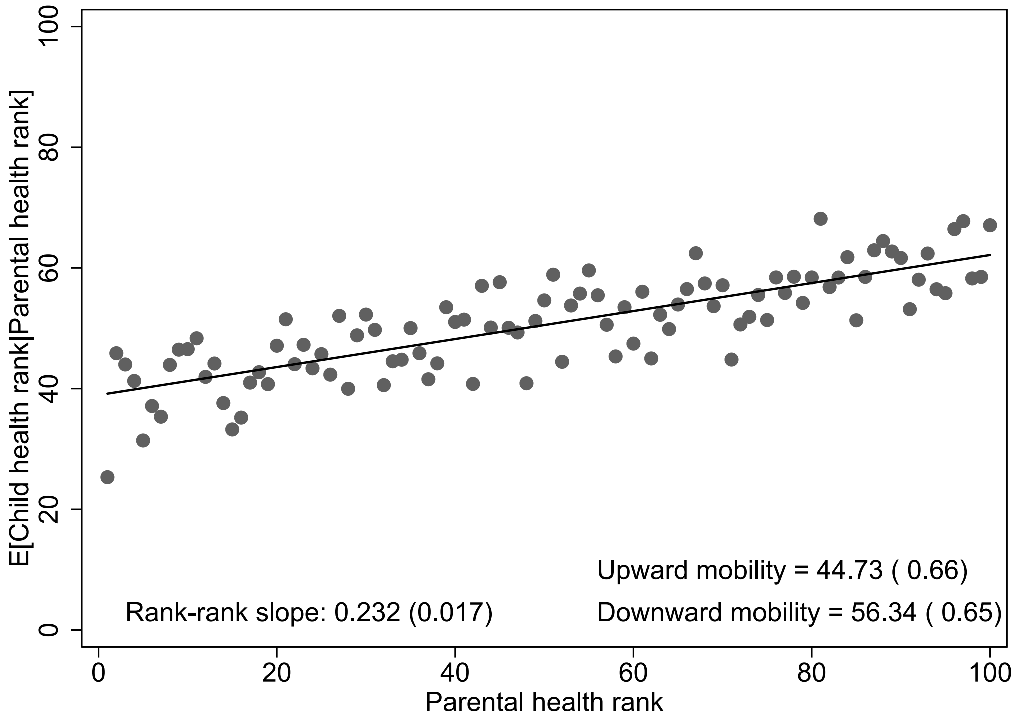
Table 1.3: Health rank mobility by parent-child relation

	Rank-rank slope (1)	Upward mobility (2)	Downward mobility (3)	Observations (4)
Mother-son	0.219	44.971	55.896	1940
	(0.023)	(0.930)	(0.824)	
Mother-daughter	0.233	44.987	56.613	1596
	(0.025)	(0.955)	(0.979)	
Father-son	0.193	45.960	55.601	1689
	(0.025)	(0.964)	(0.940)	
Father-daughter	0.198	45.866	55.758	1401
	(0.027)	(1.030)	(1.044)	
Both parents-all children	0.232	44.735	56.338	3655
	(0.017)	(0.657)	(0.647)	

Note: Each row of Table 1.3 displays the estimate of rank-rank slope, up- and downward mobility for different parent-child relations. The estimates are based on a regression of the children's percentile rank in the children's permanent health distribution on the parents' percentile rank in the parents' permanent health distribution. Robust standard errors, in parantheses, are clustered on the family level. Column (1) displays the estimates of the rank-rank slope. Columns (2) and (3) display the children's expected percentile rank if the parents' percentile rank would have been 25 and 75, respectively. Column (4) displays the number of observations.

Two patterns become apparent in Table 1.3: First, the rank-rank slopes are higher for the mother-child than for the father-child relations. This suggests a higher relative positional mobility in percentile ranks in the permanent health distribution across generations for father-child than for mother-child relations. Second, the estimates for the

Figure 1.3: Rank mobility in permanent health



Note: Figure 1.3 presents nonparametric binned scatter plots and a plot of a linear regression of the relationship between children's and parents' percentile rank in the distribution of permanent health. Each dot, corresponds to the children's average percentile rank, conditional on the parents' percentile rank. The linear fit is based on a regression of the children's percentile rank on the parents' percentile rank. Upward and downward mobility correspond to the children's expected percentile rank of parents who are located at the 25th and 75th percentile rank of the distribution of permanent health. Throughout, robust standard errors are clustered on the family level.

parent-child estimates are always higher for the daughters than for the sons, pointing to higher relative positional persistence in the distributions of permanent health across generations for daughters than for sons.[23] These observations are consistent with the findings of Halliday et al. (2021) and Andersen (2019). However, Andersen (2019) does not find any differences comparing the rank-rank slopes of father-daughter and mother-daughter relations. Notwithstanding, the rank-rank slopes suffer from ambiguous welfare implications. Therefore, we investigate the degree of up- and downward mobility in health.

We estimate an upward mobility of 44.74, as depicted in column (2) of Table 1.3. The estimate for downward mobility is 56.34, as depicted in column (3) of Table 1.3. Thus, if the parents are located at the 25th (75th) percentile, their children are

[23]However, these differences are not statistically significant.

expected to be at percentile 44.74 (56.34) in the corresponding permanent health distribution. Focusing on gender differences, we conclude that our estimates suggest that children display a higher absolute positional upward mobility in percentile ranks based on the rank-rank regressions for the father-child than for the mother-child relations. Focusing first on the measurement of upward mobility, we estimate that the degree of upward mobility is 45.87 and 45.96 for the father-daughter and father-son relations, respectively. For mothers and their children, we estimate that the upward mobility is 44.99 and 44.97 for the mother-daughter and mother-son relations, respectively. Our estimates for downward mobility are 55.76 and 55.60, for the father-daughter and father-son relations, respectively. Lastly, the estimates for downward mobility are 56.61 and 55.90 for the mother-daughter and mother-son relations, respectively.

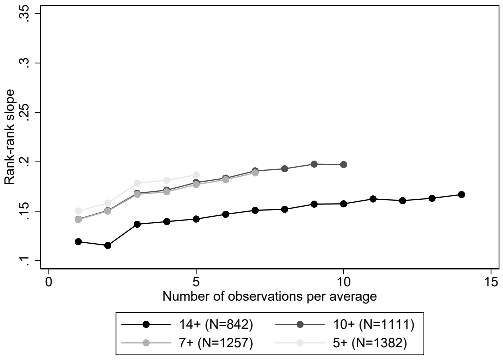
1.5 Sensitivity analysis

1.5.1 Accounting for measurement error

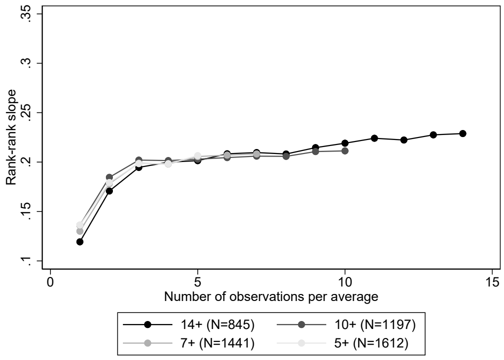
To investigate the sensitivity of our results to the presence of transitory health shocks in the explanatory variable, we construct estimates for different samples with at least z years of observations on the latent health capital per each parent. We choose z such that $z \in \{5, 7, 10, 14\}$ per parent. For each of these samples, we construct permanent health measurements based on 1 to z observations separately. Then, we perform rank-rank regressions for the z permanent health measurements in each sample. The results are displayed in Figure 1.4, in which we plot the rank-rank slope as a function of the number of observations.

Clearly, attenuation due to transitory health shocks is highly relevant. Throughout, we observe that the estimates increase with increasing numbers of observations per permanent health measurement of the parents. This confirms previous findings of Halliday et al. (2021) and Andersen (2019). However, the relevance varies across samples. Most prominently, the number of observations per average is more relevant for rank-rank regressions for sons than daughters. For sons, the estimates tend to converge if the estimates of the permanent health measurements includes at least ten observations, as displayed in Figures 1.4b and 1.4d. In contrast, the gradients are much smaller for daughters, as can be inferred from Figures 1.4a and 1.4c.

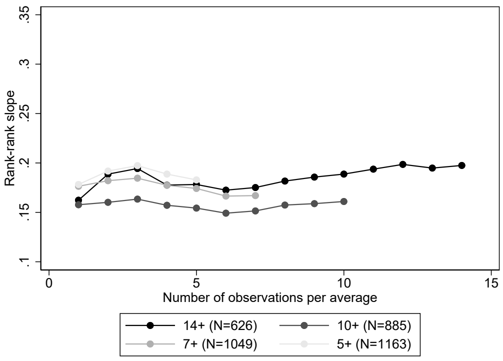
Figure 1.4: Association of rank-rank slopes and the number of observations per permanent health measurement



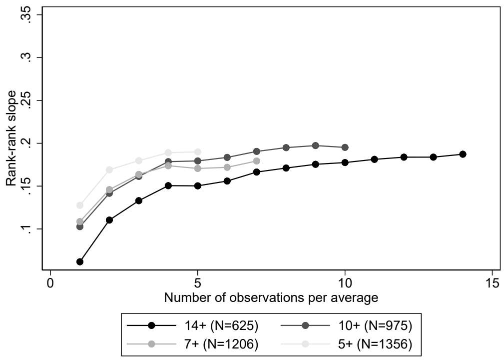
(a) Mothers and daughters



(b) Mothers and sons



(c) Fathers and daughters



(d) Fathers and sons

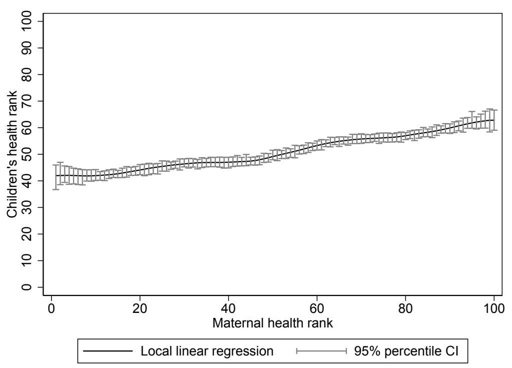
Note: Figures 1.4a to 1.4d illustrate how the rank-rank slope depends on the number of observations for the parental measurement of permanent health. The figures display the rank-rank slope of regressions of the children's percentile rank on the parents' percentile rank in the respective distribution of permanent health for different number of health observations per measurement of the permanent health per sample. The samples correspond to samples in which parents have at least 5, 7, 10 or 14 health observations available. Each figure presents results for a different parent-child sample.

A further observation is a permanent shift in the rank-rank slopes for mother-daughter and father-son samples if the sample restrictions require a higher number of observations per each parent. One possible explanation is that conditioning on the availability of at least z observations introduces unobserved heterogeneity between groups. For instance, those families that are able to contribute more observations per permanent health measurement, and which are potentially positively selected, could be hypothesized to show less persistence in health across generations. However, unfortunately, the exploration of these heterogeneities is beyond the scope of this chapter. In conclusion, one should not put to much emphasis on the levels of the rank-rank slopes in this sensitivity analysis, rather it should be on the changes as a function of the number of observations within each sample.

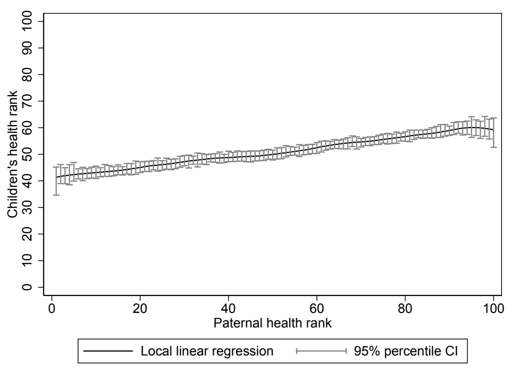
1.5.2 Linearity

In this subsection, we thoroughly investigate to what extent the assumption of linearity is warranted. Therefore, we estimate rank-rank regressions using local linear regressions. We use an Epanechnikov kernel with bandwidth w and determine the optimal bandwidth via cross-validation.[24]

Figure 1.5: Non-linear rank-rank regressions



(a) Mothers



(b) Fathers

Note: Figures 1.5a to 1.5b display the fit of a local linear rank-rank regression and the corresponding 95% percentile confidence intervals, based on bootstrapped standard errors with 50 replications each, for mothers as well as fathers and their children, respectively. We used an Epanechnikov kernel with bandwidth $w = 6$.

Figure 1.5 displays the fit of the local linear rank-rank regressions and the corresponding 95% percentile bootstrap confidence intervals, based on 50 repetitions each, for the mothers as well as fathers and their children. As it becomes immediately apparent, the estimates suggest that the association between parental and the children's percentile ranks in the respective distributions of permanent health is indeed linear. This is even more so the case for fathers and their children in contrast to mothers and their children. This is a result that is not yet shown in the literature on the intergen-

[24] We perform cross-validation for the sample of mothers and their children as well as fathers and their children separately to determine the optimal bandwidth. For mothers and their children, the optimal bandwidth is 5.79. For fathers and their children, the optimal bandwidth is significantly larger. In fact, the optimal bandwidth so large that it will always include the full support. We conclude that this is indicative of the fact that the true relation is, indeed, linear. For illustrative purposes, we impose a bandwidth of six for both samples.

erational mobility in health. This result clearly supports the linear specification of rank-rank regressions, as in Equation 1.1.

1.5.3 Life-cycle analysis

To quantify the sensitivity of our estimates to potential life-cycle biases, we consider two different age groups in both the parental and children's generation. These age groups are 30-45 and 46-65. The age range 30-45 is based on the recommendation that emerged from the literature on intergenerational income mobility (Haider and Solon, 2006; Nybom and Stuhler, 2016). Then, we test the stability of our estimates for all possible combinations of these age ranges across children and their parents.[25] The resulting estimates are displayed in Table 1.4.[26]

Overall, we find little evidence for life-cycle biases in our estimates. We find that ten out of twelve estimates are comparable to each other.[27] The two exceptions are the estimates for the sons when sons and their respective parents are 30-45, as depicted in Table 1.4. However, while these point estimates are clearly attenuated compared to those for the other sub-samples, a formal test of equality of estimates across samples does not allow us to reject the hypothesis of equality of the estimates across samples.[28] One potential explanation for the stability of the rank-rank slopes across samples could be the fact that rank-rank slopes are stable to any mean preserving spread.

[25] Since parents are older than their children and we do not have complete life-time profiles for both generations, we are not yet able to test the stability of the estimates for the sample when children are of age 46-65 and parents are of age 30-45.

[26] Please note that the estimates for the sample in which the children and the parents are of age 30-45 and 46-65, respectively, largely coincide with our main sample. Thus, large deviations between the two samples are not expected are not expected.

[27] We do not count in the estimates for the main sample in this comparison since the overlap with the sample of parents of age 46-65 and children of age 30-45 is very large.

[28] The test also includes the restriction that the estimate of the main sample are similar to those of the other subsamples. However, the null cannot be rejected in all comparisons if we exclude the latter restriction.

Table 1.4: Life cycle bias

Age group	All children-both parents	Daughter-mother	Daughter-father	Son-mother	Son-father
(1)	(2)	(3)	(4)	(5)	(6)
30-45/30-45	0.200	0.218	0.190	0.157	0.147
	(0.027)	(0.038)	(0.047)	(0.036)	(0.047)
46-65/46-65	0.250	0.208	0.182	0.246	0.216
	(0.038)	(0.059)	(0.076)	(0.055)	(0.058)
30-45/46-65	0.233	0.227	0.200	0.222	0.188
	(0.017)	(0.026)	(0.027)	(0.024)	(0.025)
30-65/30-65	0.232	0.233	0.198	0.219	0.193
	(0.017)	(0.025)	(0.027)	(0.023)	(0.025)
P-value	0.576	0.974	0.294	0.993	0.775

Note: Table 1.4 displays life-cycle patterns in the rank-rank slope for different samples. Column (1) indicates the ages under consideration for the children/parent-combination. Column (2) to (6) display the estimates for the different age combinations for the respective child-parent combination. Robust standard errors, clustered on the family level, are displayed in parantheses. The p-values in the last row are based on a Chi-square distribution, with three degrees of freedom, for a test of equality of estimates within each column. All results are based on seemingly unrelated regressions to account for potential correlation of coefficients across samples.

1.5.4 Cohort analysis

Pooling all birth cohorts, we implicitly assume that there are no differences in intergenerational mobility in health across birth cohorts of children. Indeed, our cohort specific analysis suggests that this assumption is warranted. Table 1.5 presents the rank-rank slopes and the estimates for up- and downward mobility for the different cohorts. We distinguish three birth cohorts: The first cohort is born from 1945 until 1960, the second cohort is born from 1961 to 1975, and the third cohort is born from 1976 through 1987. In addition, Table 1.5 contains p-values of tests of equality of the estimates across cohorts. Table 1.6 displays gender specific results.

Clearly, the estimates for the rank-rank slopes suggest that there is no variation in relative positional health mobility across generations. The estimates in Table 1.5 range from 0.256 for the cohort 1945-1960, to 0.238 for the cohort 1961-1975, and 0.232 for the cohort 1976-1987. Moreover, these differences are not jointly statistically significant as the p-value of 0.892 suggests. In addition, we find no differences in our measures for absolute positional up- and downward mobility in health across cohorts. The estimate for upward mobility is 44.35 for the cohort 1945-1960, 44.04 for the cohort 1961-1975, and 44.33 for the cohort 1976-1987. Further, we cannot reject the null hypothesis of no differences across estimates, as indicated by Table 1.5. The associated p-value is 0.977. Moreover, the estimates for downward mobility is 57.14 for the cohort 1945-1960, 55.94 for the cohort 1961-1975, and 55.94 for the cohort 1976-1987. Again, we find no significant differences across estimates. The corresponding p-value is 0.786, as shown in Table 1.5. Turning to gender differences, as shown in Table 1.6, we also do not detect any significant differences in health mobility across cohorts.

Table 1.5: Health mobility of children by cohort for all children

	Cohort			
	1945-1960	1961-1975	1976-1987	P-value
	(1)	(2)	(3)	(4)
Rank-rank slope	0.253	0.227	0.234	0.876
	(0.043)	(0.028)	(0.027)	
Upward mobility	44.249	44.985	44.042	0.808
	(1.641)	(1.067)	(1.051)	
Downward mobility	56.890	56.349	55.737	0.811
	(1.601)	(1.043)	(0.998)	
Observations	585	1326	1376	

Note: Table 1.5 displays the estimate of rank-rank slope, up- and downward mobility for different cohorts of children. The estimates are based on a regression of the children's percentile rank in the children's permanent health distribution on the parents' percentile rank in the parents' permanent health distribution. Upward and downward mobility are the children's expected percentile rank in the children's permanent health distribution if the parents are located at the 25th and 75th percentile rank of the parental permanent health distribution. Robust standard errors, clustered on the family level, are in parentheses. Each column corresponds to a different cohort of children. The p-values are based on a Wald Chi-square test of equality of the respective estimates based on seemingly unrelated regressions in which each cohort resembles one equation and are displayed in column (4).

Table 1.6: Health mobility of children by cohort for different subsamples

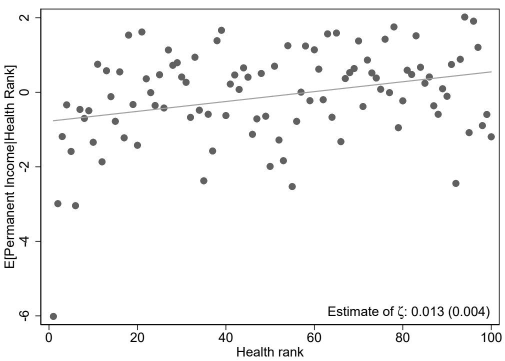
		Daughters				Sons			
		Cohort				Cohort			
		1945-1960	1961-1975	1976-1987	P-value	1945-1960	1961-1975	1976-1987	P-value
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mothers	Rank-rank slope	0.257	0.202	0.249	0.659	0.225	0.200	0.239	0.765
		(0.072)	(0.043)	(0.037)		(0.057)	(0.038)	(0.039)	
	Upward mobility	41.578	46.852	43.912	0.180	46.226	44.539	44.191	0.751
		(2.825)	(1.590)	(1.412)		(2.288)	(1.498)	(1.579)	
	Downward mobility	54.437	56.943	56.348	0.747	57.494	54.526	56.124	0.406
		(2.840)	(1.680)	(1.440)		(1.881)	(1.341)	(1.367)	
	Observations	184	552	687		354	727	665	
Fathers	Rank-rank slope	0.227	0.188	0.186	0.909	0.291	0.214	0.162	0.216
		(0.088)	(0.046)	(0.040)		(0.061)	(0.039)	(0.042)	
	Upward mobility	43.249	46.387	45.654	0.639	43.864	44.467	46.775	0.473
		(2.869)	(1.695)	(1.553)		(2.379)	(1.537)	(1.607)	
	Downward mobility	54.599	55.763	54.951	0.925	58.412	55.146	54.866	0.378
		(3.736)	(1.798)	(1.484)		(2.246)	(1.476)	(1.554)	
	Observations	141	496	629		277	632	606	

Note: Table 1.5 displays the estimate of rank-rank slope, up- and downward mobility for different cohorts of children. The estimates are based on a regression of the children's percentile rank of the children in the children's permanent health distribution on the parent's percentile rank in the parent's permanent health distribution. The p-values are based on a Wald Chi-square test of equality of the respective estimates based on seemingly unrelated regressions in which each cohort resembles one equation. Columns (1) to (4) displays estimates for daughter samples and columns (5) to (8) display estimates for son samples. Robust standard errors, in parentheses, are clustered on the family level.

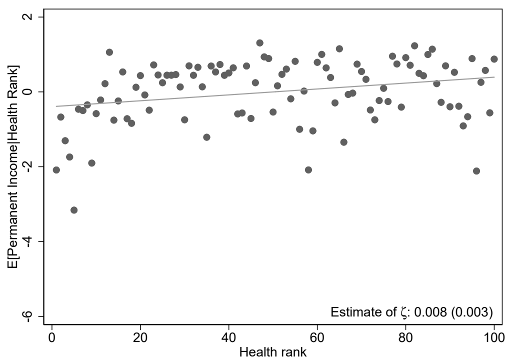
1.6 Returns to permanent health

Since our health metric exhibits no natural scale, it is impossible to evaluate movements along the health distribution. Therefore, we anchor permanent health in permanent income. To minimize potential life-cycle biases, we restrict the age range in both the children's and parents' samples to 30-45, following Haider and Solon (2006). Figure 1.6 displays the association between permanent income and the percentile rank in the distribution of permanent health for daughters, sons, mothers, and fathers.

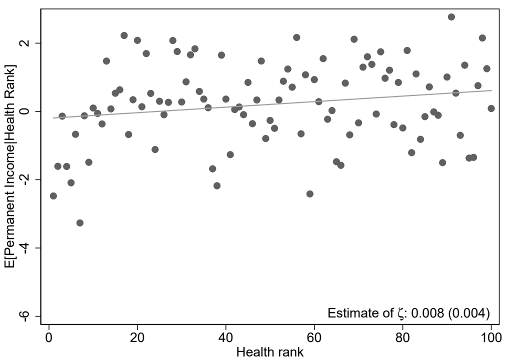
Figure 1.6: Anchoring the distribution of permanent health in permanent income



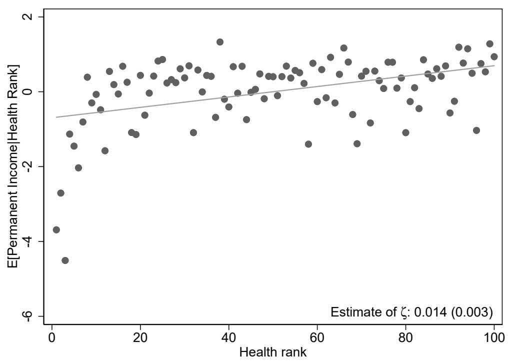
(a) Daughters



(b) Sons



(c) Mothers



(d) Fathers

Note: Figures 1.6a to 1.6d present the association between permanent income and percentile rank in the distribution of permanent health. The gray dots correspond to nonparametric binned scatter plots, displaying the sample equivalent of the population mean of the permanent income, conditional on the own percentile distribution of permanent health. The linear fit corresponds to a regression of the permanent income on own percentile rank. Robust standard errors, in parentheses, are robust to heteroscedasticity.

Clearly, we observe a non-linear and positive relation between permanent income and the percentile rank in the permanent health for the children and their parents. The non-linearity would have been masked if we had relied on a linear functional form only. Throughout, the relation appears to be linear from approximately the 20th percentile rank up to the top of the distribution of permanent health in all subsamples. In contrast, the association is stronger and highly non-linear between the bottom of the distribution of permanent health and the 20th percentile rank of the distribution of permanent health. Thus, changes in the distribution of permanent health are more consequential in the first quintile of the distribution than in other parts of the distribution of permanent health. This is consistent with Hosseini et al. (2021), who also find that it is mainly individuals scoring below the 75th percentile in their frailty index, an alternative aggregate health measure, who are more strongly affected by a change in health.

Assuming linearity, we anchor the percentile rank in the distribution of permanent health in the metric of permanent income. Consequently, we can describe the percentage change in permanent income associated with a one percentile point increase in the distribution of permanent health. A one percentile point change in the distribution of permanent health is associated with an approximate 1.3% change for daughters and a 0.8% change for sons, as inferred from Figures 1.6a and 1.6b, respectively. For parents, a one unit change in the percentile rank in the distribution of permanent health is associated with an approximate 0.8% change for mothers and 1.4% change for fathers in permanent income, respectively.

1.7 The influence of the parental SES

The strong non-linearities in the returns to health at the bottom of the health distribution points toward strong incentives for upward mobility in health. In this subsection, we investigate differences in health mobility with respect to the parents' socioeconomic background and show that an advantageous socioeconomic background is indeed associated with higher upward mobility in health. Tables 1.7 to 1.9 show the corresponding estimates for the rank-rank slopes, upward and downward mobility, as well as p-values for tests of equality of estimates across samples. The corresponding mobility curves are displayed in Section 1.E of the appendix.

Table 1.7: Health mobility by education, occupational prestige, permanent income and migration background of both parents

	Rank-rank slope (1)	Upward mobility (2)	Downward mobility (3)	N (4)
Educational degree of the parents:				
Basic or less	0.224	44.611	55.788	1165
	(0.031)	(1.056)	(1.292)	
Intermediate or more	0.221	45.211	56.279	2202
	(0.022)	(0.887)	(0.793)	
P-value test of equality	0.954	0.664	0.746	
Occupational prestige of the parents:				
Below median	0.215	45.014	55.742	1740
	(0.025)	(0.851)	(1.063)	
Above median	0.231	44.878	56.415	1627
	(0.027)	(1.125)	(0.876)	
P-value test of equality	0.658	0.923	0.625	
Permanent income of the parents:				
Below median	0.242	44.635	56.743	1745
	(0.025)	(0.844)	(1.090)	
Above median	0.201	45.691	55.739	1622
	(0.027)	(1.143)	(0.861)	
P-value test of equality	0.265	0.457	0.470	
Migration background of the parents:				
No migration background	0.244	43.733	55.926	2543
	(0.021)	(0.811)	(0.752)	
Migration background	0.185	47.787	57.028	824
	(0.035)	(1.237)	(1.547)	
P-value test of equality	0.150	0.006	0.521	

Note: Each row of Table 1.7 displays the estimate of rank-rank slope, up- and downward mobility for different subsamples, stratified according to socioeconomic characteristics of the parents. The estimates are based on a regression of the children's percentile rank in the children's permanent health distribution on the parents' percentile rank in the parents' permanent health distribution. The p-values are based on a Wald Chi-square test of equality of the rank slopes or predicted ranks across groups after a seemingly unrelated regression model, in which each subgroup corresponds to a separate equation in the seemingly unrelated regression model. Throughout, robust standard errors, in parentheses, are clustered on the family level.

Table 1.8: Health mobility of daughters by education, occupational prestige, permanent income and migration background of the mother or father

	Mother-daughter				Father-daughter			
	Rank-rank slope	Upward mobility	Downward mobility	N	Rank-rank slope	Upward mobility	Downward mobility	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Educational degree of parent:								
Basic or less	0.232	43.689	55.276	496	0.181	45.863	54.911	549
	(0.046)	(1.607)	(1.886)		(0.044)	(1.594)	(1.779)	
Intermediate or more	0.242	44.857	56.960	763	0.196	46.857	56.636	758
	(0.036)	(1.416)	(1.357)		(0.037)	(1.417)	(1.386)	
P-value test of equality	0.860	0.585	0.468		0.799	0.641	0.443	
Occupational prestige of parent:								
Below median	0.241	43.409	55.448	630	0.140	46.017	52.999	654
	(0.041)	(1.413)	(1.729)		(0.042)	(1.395)	(1.750)	
Above median	0.231	45.537	57.092	629	0.219	46.899	57.848	653
	(0.040)	(1.610)	(1.426)		(0.039)	(1.628)	(1.400)	
P-value test of equality	0.866	0.320	0.463		0.168	0.680	0.030	
Permanent income of parent:								
Below median	0.226	42.331	53.647	622	0.185	45.632	54.863	645
	(0.040)	(1.350)	(1.773)		(0.041)	(1.299)	(1.819)	
Above median	0.218	47.228	58.142	637	0.179	47.673	56.646	662
	(0.042)	(1.702)	(1.398)		(0.044)	(1.852)	(1.371)	
P-value test of equality	0.889	0.024	0.046		0.931	0.366	0.433	
Migration background of parent:								
No migration background	0.240	43.984	55.961	1029	0.192	46.111	55.715	1031
	(0.032)	(1.212)	(1.191)		(0.032)	(1.253)	(1.190)	
Migration background	0.262	45.752	58.843	230	0.203	47.289	57.447	276
	(0.066)	(2.235)	(2.885)		(0.061)	(2.000)	(2.781)	
P-value test of equality	0.759	0.486	0.354		0.872	0.617	0.566	

Note: Each row of Table 1.8 displays the estimate of rank-rank slope, up- and downward mobility for different subsamples, stratified according to background characteristics of the respective parent. The estimates are based on a regression of the percentile rank of the daughter's permanent health distribution on the percentile rank of the parent in the parent's permanent health distribution. The p-values are based on a Wald Chi-square test of equality of the rank slopes or predicted ranks across groups after a seemingly unrelated regression model, in which each subgroup corresponds to a separate equation in the seemingly unrelated regression model. Robust standard errors, in parentheses, are clustered on the family level.

Table 1.9: Health mobility of sons by education, occupational prestige, permanent income and migration background of the mother or father

	Mother-son				Father-son			
	Rank-rank slope	Upward mobility	Downward mobility	N	Rank-rank slope	Upward mobility	Downward mobility	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Educational degree of parent:								
Basic or less	0.236	42.801	54.597	647	0.218	43.890	54.785	677
	(0.040)	(1.449)	(1.557)		(0.039)	(1.345)	(1.628)	
Intermediate or more	0.180	45.517	54.495	797	0.145	48.711	55.973	869
	(0.036)	(1.489)	(1.210)		(0.036)	(1.470)	(1.221)	
P-value test of equality	0.288	0.191	0.959		0.170	0.015	0.559	
Occupational prestige of parent:								
Below median	0.227	43.938	55.281	760	0.162	46.646	54.737	810
	(0.037)	(1.387)	(1.416)		(0.037)	(1.278)	(1.492)	
Above median	0.187	44.503	53.852	684	0.209	45.601	56.073	736
	(0.038)	(1.571)	(1.298)		(0.038)	(1.604)	(1.290)	
P-value test of equality	0.449	0.788	0.456		0.368	0.610	0.498	
Permanent income of parent:								
Below median	0.225	43.314	54.566	761	0.213	46.998	57.623	788
	(0.037)	(1.328)	(1.517)		(0.038)	(1.260)	(1.611)	
Above median	0.177	45.651	54.514	683	0.183	45.005	54.150	758
	(0.039)	(1.670)	(1.225)		(0.038)	(1.633)	(1.219)	
P-value test of equality	0.372	0.273	0.979		0.581	0.333	0.085	
Migration background of parent:								
No migration background	0.240	42.568	54.581	1154	0.204	44.772	54.985	1178
	(0.029)	(1.180)	(1.030)		(0.030)	(1.169)	(1.087)	
Migration background	0.114	48.449	54.134	290	0.149	50.068	57.500	368
	(0.062)	(2.162)	(2.536)		(0.053)	(1.909)	(2.228)	
P-value test of equality	0.062	0.017	0.870		0.360	0.018	0.309	

Note: Each row of Table 1.9 displays the estimate of rank-rank slope, up- and downward mobility for different subsamples, stratified according to background characteristics of the respective parent. The estimates are based on a regression of the percentile rank of the daughter's permanent health distribution on the percentile rank of the parent in the parent's permanent health distribution. The p-values are based on a Wald Chi-square test of equality of the rank slopes or predicted ranks across groups after a seemingly unrelated regression model, in which each subgroup corresponds to a separate equation in the seemingly unrelated regression model. Robust standard errors, in parentheses, are clustered on the family level.

1.7.1 Income

An influential literature investigates the gradient in the association between parental income and children's health (e.g. Blau, 1999; Case et al., 2002; Reinhold and Jürges, 2012; Khanam et al., 2009; Currie et al., 2007; Currie and Stabile, 2003; Apouey and Geoffard, 2013; Kuehnle, 2014). Typically, parents with higher income are hypothesized as being able to provide a more favorable environment for their children, e.g. they can buy healthier food, medical services, or are better able to adhere to medical instructions (Case et al., 2002). Three stylized facts emerged in this literature: First, the gradient between parental income and children's health increases with the children's age. Second, the association between parental income and children's health is mainly attributed to differences in the severity of health conditions in contrast to the prevalence of health conditions. Third, it is permanent income rather than contemporaneous income that matters. In what follows, we compare intergenerational mobility in health for children of parents with permanent income above and below the median.

Table 1.7 shows no difference between children of parents with high and low permanent income. However, we find that the daughters' health mobility depends on mother's permanent income. Table 1.8 shows that daughters of mothers with high income have better health throughout. The differences of the estimates in upward and downward mobility are 4.90 and 4.50 and significant in a SUR framework, as the p-value in Table 1.8 indicates. This is a remarkable result. It suggests that the distribution of permanent health of daughters of mothers with permanent income above the median first-order stochastically dominates the distribution of permanent health of daughters of mothers with permanent income below the median.[29] Thus, under reasonable assumptions about the individuals' preferences, daughters would prefer to be born into families in which mothers have permanent income above the median. This is a result that would have been masked if we would have only focused on the rank-rank slope or intergenerational health associations. The rank-rank slope of daughters

[29] If we compare two cumulative distribution functions, e.g. $F_a(x)$ and $F_b(x)$, and suppose x is a desirable outcome, such as health, then we say that option a first-order stochastically dominates b if $F_a(x) \leq F_b(x)$. An immediate consequence is that for any utility function $u(\cdot)$, we have that $\int u(x)dF_a(x) \geq \int u(x)dF_b(x)$. To put it differently, every utility maximizing individual should choose option a over b since option a maximizes the expected utility of this individual.

of mothers with high and low permanent income are 0.226 and 0.218, respectively. The difference of these estimates is small and not statistically significant, as Table 1.8 shows.

1.7.2 Education

A burgeoning literature investigates the effect of parental education on children's health (Thomas et al., 1991; Lindeboom et al., 2009; Lundborg et al., 2014; Silles, 2015; Kemptner and Marcus, 2013; Huebener, 2018). In this literature, parental education is hypothesized to affect the children's health via improved parental behavior or increased financial resources (e.g. Lindeboom et al., 2009). While still inconclusive, the majority of studies in this literature points toward a positive effect of parental education on children's health. In our analysis, we distinguish children of parents who have at least an intermediate school leaving degree from parents who have at most a basic school leaving degree. Table 1.7 displays no difference between children of parents with a high and low school leaving degree. Further, if we additionally distinguish between genders, we conclude that sons of fathers with higher education experience a greater upward mobility in health than sons of fathers with lower education. The difference in upward mobility amounts to 4.92 and is statistically significant different from zero, as Table 1.9 suggests.[30]

1.7.3 Occupational prestige

A low occupational status is often associated with physical and mental strain as well as low job control (Ravesteijn et al., 2018). In particular, the stress associated with low occupational status can negatively affect the interaction between parents and their children. Moreover, the literature shows that stress on the parents' side limits the attention parents can give to parenting (e.g. Cobb-Clark et al., 2019). Consequently, we expect health mobility to differ by parental occupational prestige. For this, we distinguish between children whose parents' occupational prestige score is below and

[30] At first glance, this contrasts with evidence by Huebener (2018) who reports no effect of paternal years of schooling on their children. However, this could be explained by the fact that Huebener (2018) exploits a reform that increased the number of compulsory years of schooling from eight to nine years. Formally, these parents were still entitled to a basic school leaving degree only. Thus, the corresponding local average treatment effect applies to individuals at the lower end of the educational hierarchy. In contrast, we compare children of parents with different school leaving degrees.

above the median. Indeed, daughters of fathers with occupational prestige above the median are indeed more upwardly mobile, as Table 1.8 shows. The difference in percentile ranks is 4.85 and statistically significantly different from zero.

1.7.4 Migration background

In social sciences, the literature documents a "healthy immigrant effect" (e.g. Antecol and Bedard, 2006; Domnich et al., 2012; Jasso et al., 2004; zur Nieden and Sommer, 2016; Palloni and Arias, 2004; Razum et al., 1998; Ullmann et al., 2011; Giuntella and Mazzonna, 2015; Kennedy et al., 2006; Antman et al., 2020). The "healthy immigrant effect" describes the empirical phenomenon that immigrants are healthier than the native population. Typical explanations for this are selection, health behaviors, and return migration (e.g. Giuntella and Mazzonna, 2015). In addition, a scarce literature focuses on health differences between children of immigrant and native born parents (e.g. García-Pérez, 2016; Kotwal, 2010; Razum et al., 1998). In the U.S., evidence points toward a convergence, if not reversal, of the initial health advantage of immigrants across generations (García-Pérez, 2016). For Germany, epidemiological studies suggest a persistence of the "healthy immigrant effect" into the second generation (Kotwal, 2010; Razum et al., 1998). We contribute to this evidence by showing that this health advantage is driven by sons stemming from parents at the lower end of the health distribution. In our analysis, we compare children of parents who are and who are not born in Germany. The difference amounts to 5.05 and is statistically significant in a SUR framework, as presented in Table 1.7. A gender specific analysis reveals that this is mainly driven by sons and their parents. Table 1.9 shows that the difference in upward mobility is 5.88 for the mother-son and 5.30 for the father-son relation.

1.8 Conclusion

Surprisingly, very little evidence exists on intergenerational mobility in health, despite the fact that health is a crucial determinant of individuals' (economic) well-being. We provide the first estimates of intergenerational mobility in permanent health for Germany using 25 years of detailed health information for two generations. For this, we model the joint distribution of permanent health for parents and their children. Thus, we estimate rank-rank regressions, regressing the percentile rank of the children on the percentile rank of the parents in their respective distributions of permanent health. In our main results, we find a rank-rank slope of 0.232 and find that up- and downward mobility in permanent health are 44.74 and 56.34, respectively. In addition, we find a positive and non-linear association between permanent health and income, providing incentives to escape the bottom of the distribution of permanent health. And indeed, we find that a more favorable parental SES is associated with upward mobility in permanent health across generations.

A naturally arising question is: How does health mobility inform us about the state of societies? Good health is the precondition for individuals to exert any effort aiming at increasing income and consumption possibilities (Sen, 2002). Therefore, one could certainly consider intergenerational health mobility a characteristic of an egalitarian society. For earnings, a mobile society can additionally be considered a society that rewards effort. This also applies to health to the extent that good health increases individuals' productivity. However, while most studies acknowledge a role for productivity effects of health, most studies conclude that it is the employment channel that drives the association between earnings and health (e.g. Britton and French, 2020; French, 2005; Hosseini et al., 2021). This aligns more closely with the notion that bad health limits individuals' capabilities to exert any effort. Therefore, we argue that health mobility is more a sign of an egalitarian society rather than a sign of a society that rewards effort.

Appendix

1.A Details on the IRT model

In this section, we provide details on the IRT model. Let $A_{ij} = k$ correspond to the answer to item j given by individual i, with $k \in M_j = \{m_1, m_2, \dots, m_{h_j}\}$, such that $m_1 < m_2 < \dots < m_{h_j}$. M_j is of cardinality h_j. The scalar $H \in \mathbb{N}$ is the number of health items. The GRM explicitly models the probability of observing answer m_k or higher for item j for individual i as a non-linear function of the latent health status θ_i. Thus, in a first step, we estimate the parameter space $\mathbf{B} = \{\mathbf{a}, \mathbf{b} \mid \mathbf{a} \in \mathbb{R}^H, \mathbf{b} \in \mathbb{R}^P\}$, in which $\mathbf{a}$ corresponds to vector $(\alpha_1, \dots, \alpha_H)$ and $\mathbf{b}$ to vector $(b_{11}, b_{12}, \dots, b_{1h_1}, b_{21}, \dots, b_{Hh_H})$, and P is the sum over all h_j, of the model

$$P(A_{ij} \geq m_k \mid \theta_i) = \frac{\alpha_j exp(\theta_i - b_{jk})}{1 + \alpha_j exp(\theta_i - b_{jk})}. \tag{1.4}$$

In Equation 1.4, parameter α_j describes the discriminatory power of item j and b_{jk} is the difficulty parameter associated with each potential response for each item. The probability of observing answer k to item j for each respondent i is then given by the empirical mean of

$$P(A_{ij} = m_k \mid \theta_i) = P(A_{ij} \geq m_k \mid \theta_i) - P(A_{ij} \geq m_{k+1} \mid \theta_i). \tag{1.5}$$

After estimating the parameters in $\mathbf{a}$ and $\mathbf{b}$, we estimate the value of the latent health status θ_i in a second step by the empirical Bayes method. Assuming that $\theta_i \sim \mathcal{N}(0, 1)$, it follows that

$$\hat{\theta}_i = \int_{-\infty}^{\infty} \theta_i \prod_{j=1}^{H} \frac{P(A_{ij} = k \mid \theta_i, \hat{\alpha}_i, \hat{b}_{jk}) f(\theta_i)}{\int_{-\infty}^{\infty} \prod_{j=1}^{H} P(A_{ij} = k \mid \theta_i, \hat{\alpha}_i, \hat{b}_{jk}) f(\theta_i) d\theta_i} d\theta_i. \tag{1.6}$$

The resulting estimate $\hat{\theta}_i$ proxies the contemporaneous latent health status or health capital of the individual i.

1.B Life-cycle profile of the latent health status

The literature on intergenerational income mobility emphasizes that is of utmost importance to take earnings observations from an age range between the early thirties and mid-forties (Haider and Solon, 2006; Nybom and Stuhler, 2016, 2017). Otherwise, earnings observations or averages of multiple earnings observations are likely to be a biased proxy for permanent income or life-time earnings. The sources of these biases are heterogenous earnings growth rates over the life-cycle across individuals. For instance, individuals with higher permanent income typically have lower earnings at the beginning of their life but steeper earnings growth rates later in the life-cycle than individuals with lower permanent income. One reason for these heterogenous growth rates are different propensities to invest in human capital (Mincer, 1958; Ben-Porath, 1967; Becker, 1962).

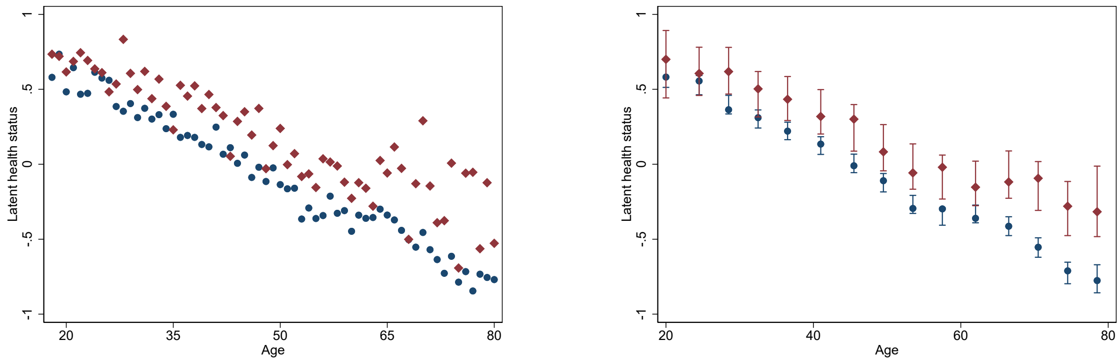
Similar to earnings, health is also potentially prone to heterogeneous changes over the life-cycle across individuals with different SES. Therefore, we illustrate the life-cycle pattern in health in this section. To be precise, we lay out some life-cycle facts of our health measure, focusing specifically on a single cross-section in year 2007. We restrict the age range to observations between age 18 and 80. Next, we calculate the mean for each age-paternal education-cell. The resulting plot is displayed in Figure 1.B.1a. Figure 1.B.1b corresponds to the binned scatter plot based on the data for Figure 1.B.1a.[31] Figure 1.B.1a and 1.B.1b together lead to our first observation:

1. *Differences that are traced back to the parental background stay latent until mid-life and become salient thereafter.*

Clearly, the averages of both groups overlap largely in Figure 1.B.1a. Further, the confidence intervals in Figure 1.B.1b suggest that the health status is statistically indistinguishable between groups until the age of thirty. Thus, like earnings, no clear ordering of the two groups with respect to health can be established until the age of approximately 30. After the age of 30, a clear ordering emerges, with rather stable differences between groups until the retirement age. Thus, relying on observations before age 30 could result in misleading conclusions. This is similar to the problem

[31] We use the Stata package **binsreg**. For the documentation, please refer to Cattaneo et al. (2019).

Figure 1.B.1: Life-cycle profile of the latent health status



(a) Raw data

(b) Binscatter

Note: Figures 1.B.1a to 1.B.1b display the life-cycle trajectories in latent health by age and paternal education. Figure 1.B.1a displays the average latent health status by age and education level. Figure 1.B.1b displays binned scatter plots with 15 bins. In Figure 1.B.1b, the dots correspond to means of the bins. The vertical bars correspond to 95% confidence intervals, based on robust standard errors for polynomials of degree three with three smoothness constraints. Blue dots correspond to the averages of children whose fathers have no or a basic school leaving degree. Red dots correspond to the averages of individuals whose parents have a tertiary school leaving degree.

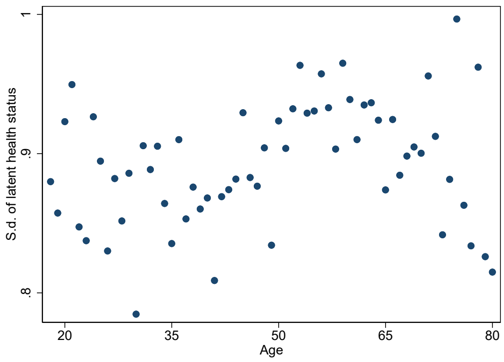
with early observations for earnings, when orderings are not well established. This is why we take observations starting from 30 until retirement.

Moreover, in Figure 1.B.2a and 1.B.2b, we display the standard deviation in latent health by age. Clearly, the standard deviation in latent health increases with age. Therefore, we summarize the following:

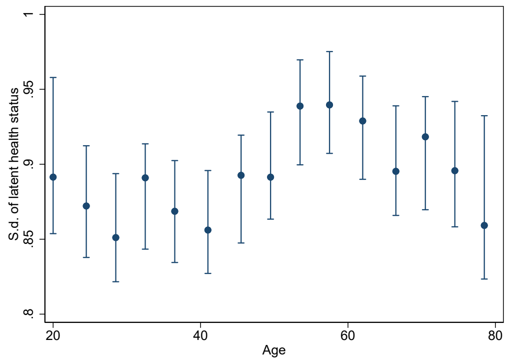
2. *Health inequalities increase with age.*

This is consistent with previous evidence on health inequalities over the life-cycle (Timothy, 2011; Deaton and Paxson, 1998; Halliday et al., 2020), resulting in important implications for the measurement of intergenerational persistencies in health. If one regresses children's on parents' health, results will clearly depend on the age range at which parents' or children's health is measured. If health inequalities increase with age, as it is displayed in Figure 1.B.2a or 1.B.2b, health associations decrease as we take parental observations from older than younger ages, *ceteris paribus*. Similarly, OLS associations would increase as we take observations from higher ages than lower ages for children. This is a direct consequence of the observation that the OLS coefficient corresponds to the linear correlation between the two outcomes, rescaled by the ratio of the standard deviation of the children's and parents' health outcomes. This

Figure 1.B.2: Inequality in health over the life-cycle



(a) Raw data



(b) Binscatter

Note: Figures 1.B.2a to 1.B.2b display the standard deviation in the latent health status over age. Figure 1.B.2a displays the standard deviation of the latent health status by age. Figure 1.B.2b displays the corresponding binned scatter plot with 15 bins. In Figure 1.B.2b, the dots correspond to means of the respective ages. The vertical bars correspond to 95% confidence intervals, based on robust standard errors for polynomials of degree three with three smoothness constraints.

is not the case for rank-rank regressions, which are invariant to any mean preserving spread. This is a direct consequence of the fact that, for the uniform distribution, with lower and upper bound equal to zero and one, the variance is equal to $\frac{1}{12}$. Thus, changes in the variance with age are not relevant for rank-rank regressions.

1.C Detailed information on the utilized health information in the SOEP and the calibration of the IRT model

In this section, we provide more detail on the health items as well as the calibration of the IRT model. The self-rated health status is inferred by the answer to the question "How would you describe your current health?" Answers are given on a five point Likert-scale ranging from one "Very good" to five "Bad." The self-rated health status is shown to be highly predictive for illnesses and mortality, even after conditioning on objective health information (see van Doorslaer and Gerdtham (2003), Pijoan-Mas and Ríos-Rull (2014), Miilunpalo et al. (1997) for an overview; Schwarze et al. (2000) provide evidence for the SOEP).

A related subjective measure of the current health status is satisfaction with health. Satisfaction with health is inferred by the answer to the question "How satisfied are you with your health?" Answers are given on an eleven point Likert-scale ranging from zero "Completely dissatisfied" to ten "completely satisfied."[32]

The two reported objective measures for in- and outpatient care are doctor visits within the last three months as well as hospital admissions in the previous year. Noteworthy, we dichotomize the number of hospital visits and the number of doctor visits within three months prior to the interview. The reason is that we want to ensure that the items are reflective of health and not a formative factor of health. One could, for instance, argue that preventive care can be a cause of good health. Thus, for hospital visits, we construct an indicator that is equal to one if an individual has been admitted to the hospital at least twice in the previous year.[33] For doctor visits, we construct an indicator that is equal to one if an individual visited the doctor more than three times in the last three months.

The objective measure for health is the degree of disability or reduced earnings capacity. In Germany, individuals can apply to have their disability ascertained by a medical reviewer. The degree of disability is then documented, allowing the individual

[32] We reversed the scale such that all scales of the health items have the same polarization.

[33] We use two hospital visits since giving birth is associated with hospital stays for females. We argue that this is not necessarily related to health.

to access compensation, including tax allowances, additional vacation days, and early retirement, among others. The process is highly formalized and documented in a bylaw enacted by the federal government.[34] The degree of disability starts at 0, indicating the absence of disabilities, and ranges to 100 in increments of 10. Additionally, the reduced earnings capacity captures the degree to which individuals are incapacitated for work. Again, this is highly regulated and exact formulations are found in the social security code[35]. Realizations of the degree of reduced earnings capacities can potentially range from 0 to 100.[36] We discretize the degree of disability into eleven categories.

We calibrate the IRT model for the full population of the SOEP in 2006, the middle of the observation period, and for all respondents providing answers to all items. [37] However, in a first step, we show that the correlation between the five health items is consistent with a unique trait causing the co-movement of the health items.

A principal component analysis of the health items for the population of the SOEP in 2006 shows that the items load unambiguously on one factor. Based on that, we conclude that the items are reflective of an underlying trait, which we refer to as latent health status. For instance, Figure 1.C.1 plots the factors and their corresponding Eigenvalues in descending order according to the magnitude of the respective Eigenvalue. While the first factor has an Eigenvalue of 2.37, the second factor has an Eigenvalue of 0.94, which is below the threshold of 1 implied by the Kaiser criterion (Kaiser and Dickman, 1959). Lastly, the second factor is the factor at which the curve in the screeplot levels off, leaving the first factor as the only significant factor (Cattell, 1966).[38] Further, the factor loadings, depicted in Table 1.C.1, range from 0.87 for the self-rated health status to 0.36 for the hospital visits in the previous year. Three out of five items are associated with factor loadings of 0.85 or higher. Throughout, the self-reported health measures and the degree of disability are associated with higher factor loadings than the reports of out- and inpatient care.

[34] "Verordnung zur Durchfuehrung des §1 Abs. 1 und 3, des §30 Abs. 1 und des §35 Abs. 1 des Bundesversorgungsgesetzes (Versorgungsmedizin-Verordnung - VersMedV)"

[35] "Sozialgesetzbuch 6, §43"

[36] The degree of disability and the degree of reduced earnings capacity are assessed within the same item in the SOEP.

[37] This procedure is similar to the routine to calibrate the physical and mental scale of the SF12v2 in the SOEP (Nuebling et al., 2007).

[38] This is the intuition of the "Elbow-criterion."

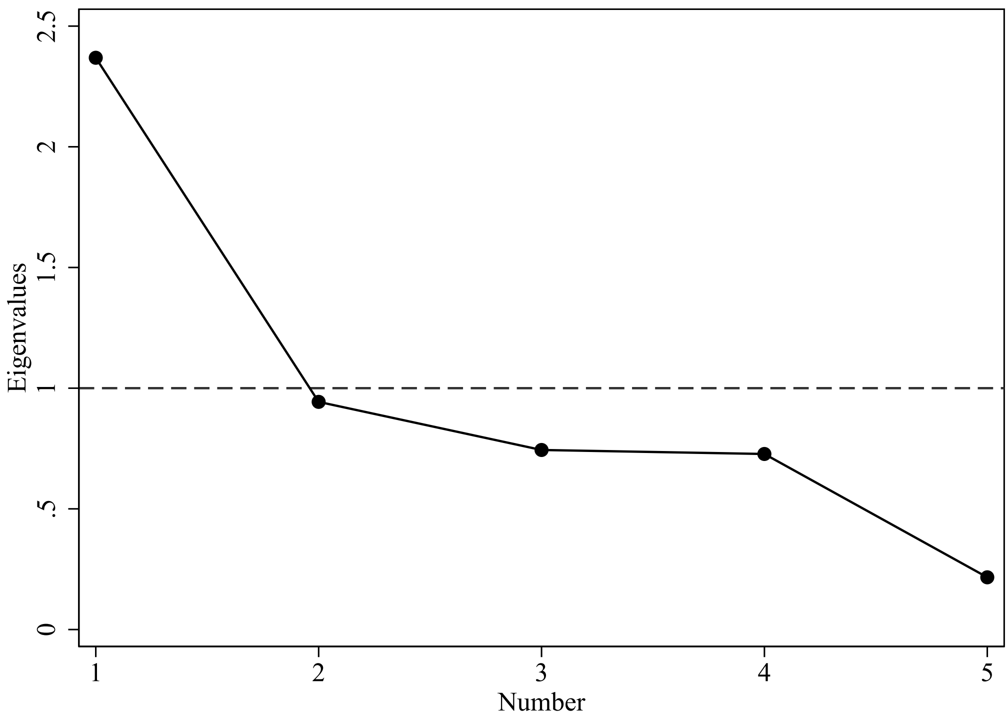
Table 1.C.1: Factor loadings of a principal component analysis of health items of the SOEP population in the survey year 2006

Item (1)	Factor loading (2)
Self-rated health status	0.87
Satisfaction with health	0.85
More than three doctor visits in last three months	0.61
More than one hospital visit in previous year	0.36
Degree of disability	0.87

Note: Table 1.C.1 displays the factor loadings of a principal component analysis of the recoded health items in the SOEP survey year 2006. The sample is restricted to full response on the five health items. Column (1) shows the recoded health item. Column (2) displays the corresponding factor loadings for the first factor.

In a second step, we calibrate the GRM model and predict the latent health status for all individuals in all years for which we observe full item response on the five health items. In a third step, we keep all individuals in the age range 30 to 65 in the children's and parent's generation. Then, in a last step, we take the individual time average using all available observations for individuals, after accounting carefully for age as well as year of birth fixed effects. The last step accounts for transitory shocks to health. Otherwise, our estimates may suffer from attenuation bias, as shown in the sensitivity analysis.

Figure 1.C.1: Screeplot of principal component analysis of health items in 2006



Note: Figure 1.C.1 plots the Eigenvalue of a principal component analysis of self-rated health status, satisfaction with health, an indicator for having visited the doctor more than three time in the last three months, an indicator for being admitted to the hospital for at least two times in the previous year, and a discretized version of the degree of disability against the corresponding factors, with the factors being ordered in descending order. The dashed horizontal line indicates Eigenvalues with a value of one.

1.D Additional tables

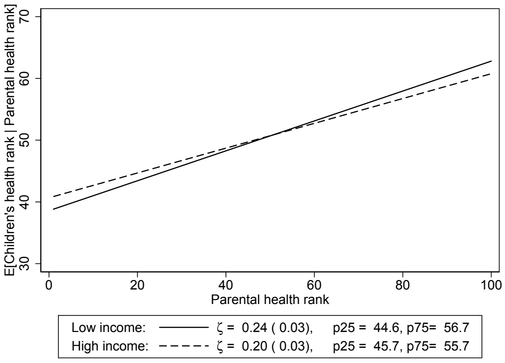
Table 1.D.1: Health items for calculation of latent health status

Item (1)	Wording of question (2)	Potential answers (3)	Recoding (4)
Self-rated health status	"How would you describe your current health?"	5-point Likert-scale ranging from 1 "Very good" to 5 "Bad"	None
Satisfaction with health	"How satisfied are you with your health?"	11-point Likert-scale ranging from 0 "completely dissatisfied" to 10 "completely satisfied"	Scale reversed so that higher values indicate worse health
Doctor visits last 3 months	"Have you gone to a doctor within the last three months? If yes, please state how often."	Open answer	Indicator for having visited the doctor more than three times in the last three months
Hospital visits previous year	"What about hospital stays in the last year - were you admitted to a hospital for at least one night in $t-1$?"	Open answer	Indicator for being admitted at least twice in the previous year
Degree of disability	"Have you been officially assessed as being severely disabled or partially incapable of work for medical reasons?"	Open answer	Discretized to eleven values ranging from 0 to 100 in increments of 10.

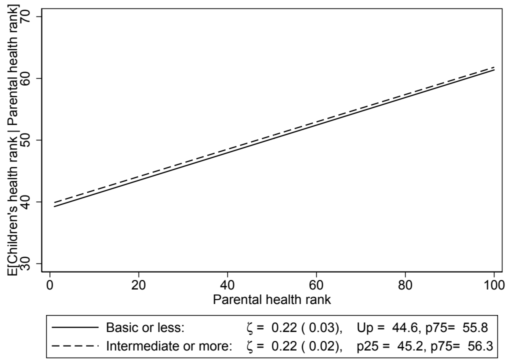
Note: Table 1.D.1 displays the health items in the SOEP used to predict the latent health status. Column (1) displays the title of the health item. Columns (2) and (3) display the exact wording of the item as well as the potential answers. Column (4) displays the recoding of the health item.

1.E Additional figures

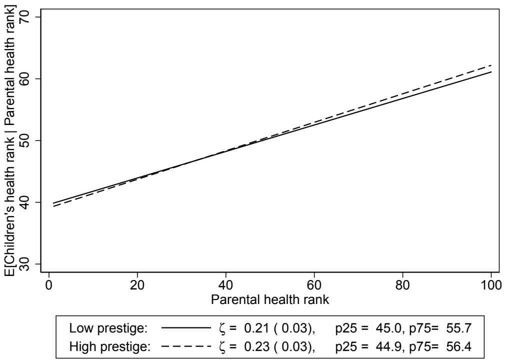
Figure 1.E.1: Parental socioeconomic characteristics and health mobility



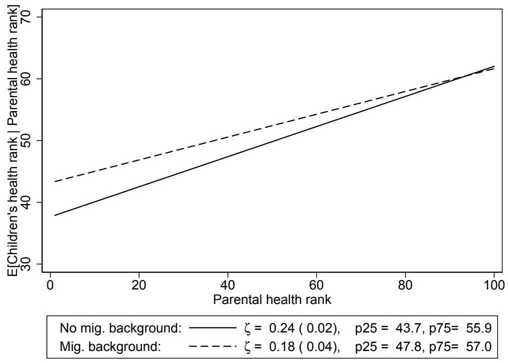
(a) Permanent income



(b) Education



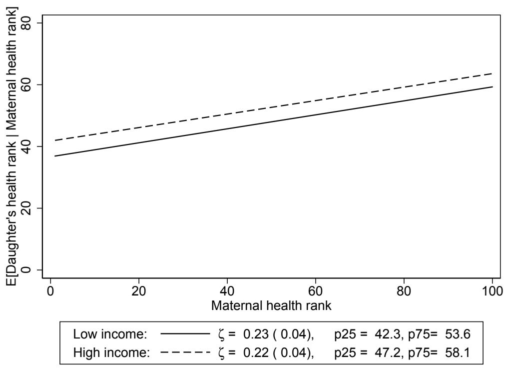
(c) Occupational prestige



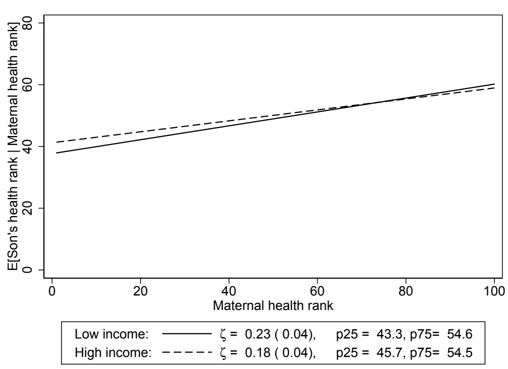
(d) Migration background

Note: Figures 1.E.1a to 1.E.1d display the variation in intergenerational mobility in health with parental socioeconomic characteristics. Each figure displays a linear fit of a regression of the children's percentile rank on the parents' percentile rank for subgroups. The respective estimates of up- and downward mobility are denoted p25 and p75, respectively. Robust standard errors are clustered on the family level.

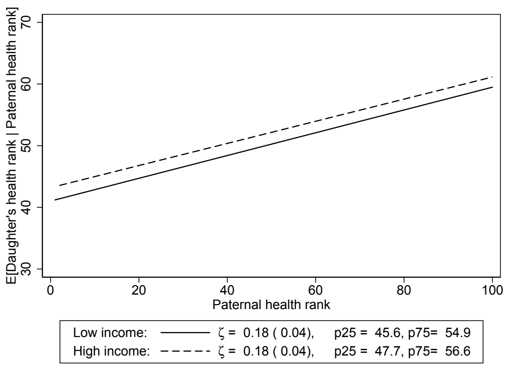
Figure 1.E.2: Parental permanent income and health mobility



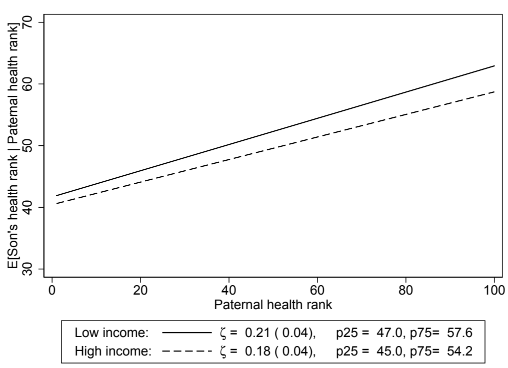
(a) Mothers and daughters



(b) Mothers and sons



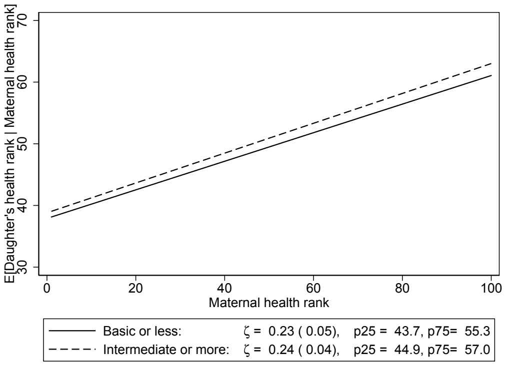
(c) Fathers and daughters



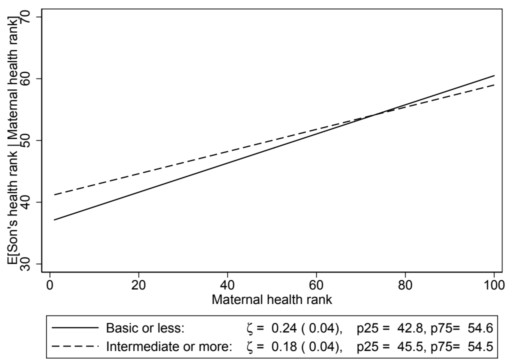
(d) Fathers and sons

Note: Figures 1.E.2a to 1.E.2d display the variation in intergenerational mobility in health with respect to parental permanent income for different parent-child relations. Each figure displays a linear fit of a regression of the children's percentile rank on the parents' percentile for parents with high and low permanent income. The respective estimates of up- and downward mobility are denoted p25 and p75. Robust standard errors are clustered on the family level.

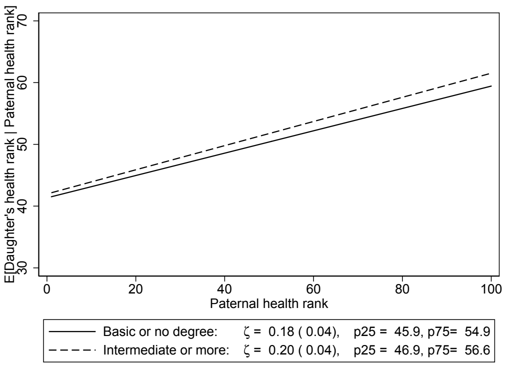
Figure 1.E.3: Parental education and health mobility



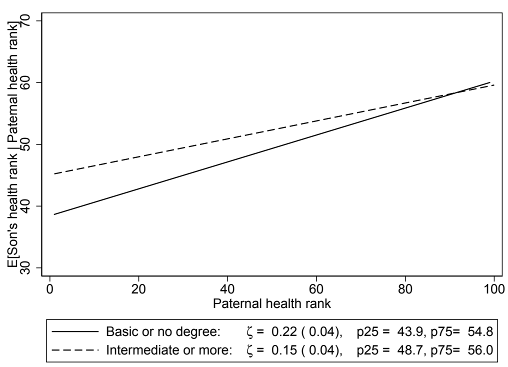
(a) Mothers and daughters



(b) Mothers and sons



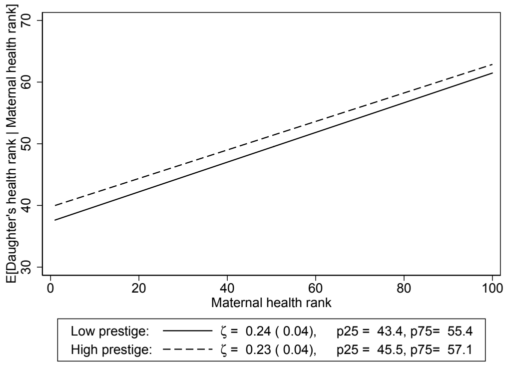
(c) Fathers and daughters



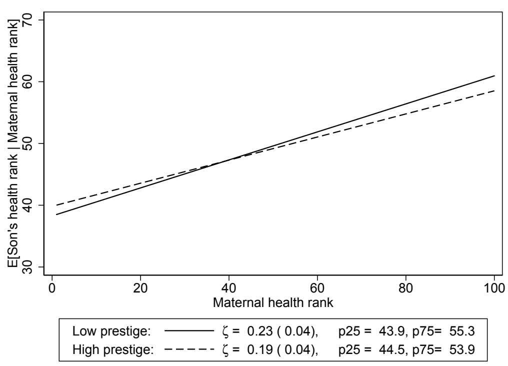
(d) Fathers and sons

Note: Figures 1.E.3a to 1.E.3d display the variation in intergenerational mobility in health with respect to the parental education for different parent-child relations. Each figure displays a linear fit of a regression of the children's percentile rank on the parents' percentile rank for parents with low and high education. The respective estimates of up- and downward mobility are denoted p25 and p75. Robust standard errors are clustered on the family level.

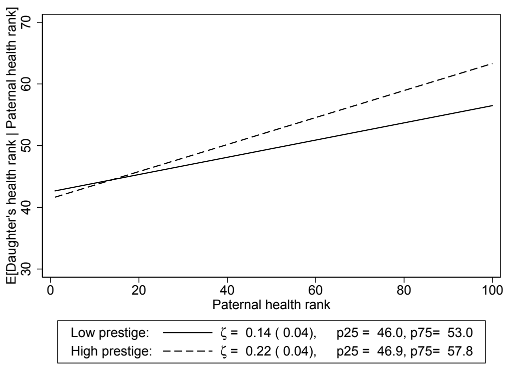
Figure 1.E.4: Parental occupational prestige and health mobility



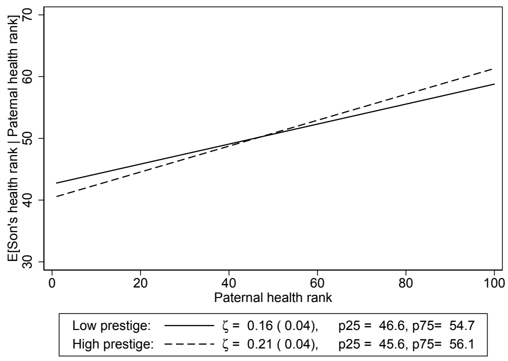
(a) Mothers and daughters



(b) Mothers and sons



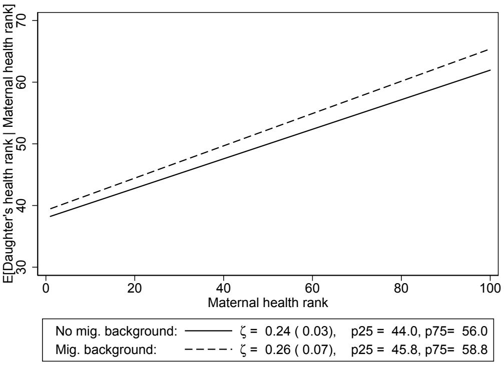
(c) Fathers and daughters



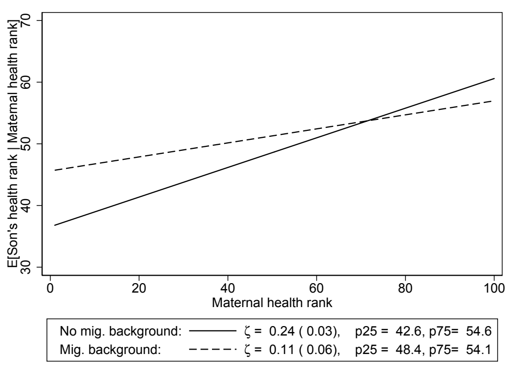
(d) Fathers and sons

Note: Figures 1.E.4a to 1.E.4d display the variation in intergenerational mobility in health with respect to the occupational prestige for different parent-child relations. Each figure displays a linear fit of a regression of the children's percentile rank on the parents' percentile rank for parents with high and low occupational prestige. The respective estimates of up- and downward mobility are denoted p25 and p75. Robust standard errors are clustered on the family level.

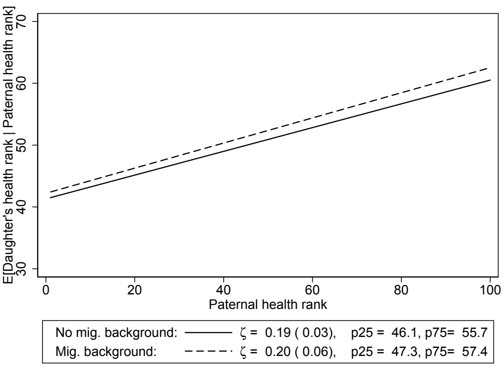
Figure 1.E.5: Parental migration background and health mobility



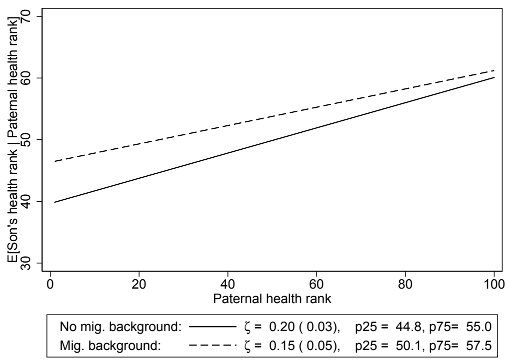
(a) Mothers and daughters



(b) Mothers and sons



(c) Fathers and daughters



(d) Fathers and sons

Note: Figures 1.E.5a to 1.E.5d display the variation in intergenerational mobility in health with respect to the parental migration background for different parent-child relations. Each figure displays a linear fit of a regression of the children's percentile rank on the parents' percentile rank for parents with and without migration background. The respective estimates of up- and downward mobility are denoted p25 and p75, respectively. Robust standard errors are clustered on the family level.

CHAPTER 2

The effect of maternal education on offspring's mental health*

We estimate the effect of maternal schooling on children's mental health in adulthood. Using the Socio-Economic Panel and the mental health measure based on the SF-12 questionnaire, we exploit a compulsory schooling law reform to identify the causal effect of maternal schooling on children's mental health. While the theoretical considerations are not clear, we do not find that the mother's schooling has an effect on the mental health of the children. However, we find a positive effect on children's physical health operating mainly through physical functioning. In addition, albeit with the absence of a reduced-form effect on mental health, we find evidence that the number of friends moderates the relationship between maternal schooling and their children's mental health.

*This chapter is joint work with Daniel Schnitzlein. A previous version has been published as "The effect of maternal education on offspring's mental health." SOEPpapers on Multidisciplinary Panel Data Research 1028. This version has been submitted to the *Journal of Behavior & Organization.*

2.1 Introduction

Worldwide, mental health conditions are a leading cause of disability-adjusted life years (DALYs) and increasing health costs. They account for 199 million DALYs or 37% of healthy life years lost due to noncommunicable diseases. The sum of direct and indirect costs in 2010 is estimated to equal 2.5 trillion US dollars and projected to increase to 6 trillion US dollars in 2030 (Bloom et al., 2011). Given that the financial and societal burdens of mental health impairments are high, prevention measures that alleviate mental health problems will have high financial and societal returns. In addition, contributions have shown a strong intergenerational transmission of mental health status (Johnston et al., 2013) implying long-run consequences for the children of those affected by mental health problems today.

This chapter is the first to estimate the effect of maternal education on children's mental health in adulthood. Given the high prevalence rate of mental health issues, easing the burden of mental health problems would have immediate payoffs. For instance, Layard (2016) estimates that a relief of societies' burden of mental illnesses could increase general employment by four percent and thus increase GDP. Therefore, our results are highly relevant for policy makers who are eager to alleviate the burden of mental health problems.

Theoretical considerations are ambiguous about the effect of maternal education on children's mental health. Moreover, the empirical relation between maternal education and children's mental health is potentially subject to endogeneity. For this reason, we exploit exogenous variation caused by a compulsory schooling law (CSL) reform that extended compulsory schooling from eight to nine years across states and time in West Germany. Since it is only mothers at the lower end of the educational hierarchy for whom this CSL reform is binding, our study provides empirical evidence regarding how to alleviate the socioeconomic gradient in health across generations (Currie, 2009; Case et al., 2005; Ahlburg, 1998, e.g.).

Estimating the intergenerational effect of maternal education on the mental health of children is a difficult task. On the one hand, we need information about parents and their children. On the other hand, we need information about the children's mental health. Information about both at the same time is rarely available in datasets. The

Socio-Economic Panel (SOEP) entails extensive information about parent-child pairs, which makes it especially well-suited for our study. Our principal measure of mental health is the Mental Component Summary (MCS) score, based on the Short-Form 12 (SF-12) questionnaire, which comprises twelve health-related questions covering both mental and physical health dimensions. The MCS score is a well-established measure for mental health in the epidemiological literature and is shown to be predictive for mental illnesses (Salyers et al., 2000). Another advantage of the SF-12 questionnaire is that it allows us to directly compare the effect of the CSL reform on mental health with the effect on physical health. Along with the MCS score, the SF-12 allows deriving the Physical Component Summary (PCS) score. The PCS score is the equivalent of the MCS score for the physical health dimension.

A large and active literature investigates the causal effect of parental education on their children's health. Nevertheless, the investigation of the effect of parental education on children's health is complicated by three aspects: first, unobserved characteristics that are associated with education and mental health and that are transmitted across generations are likely to confound OLS estimates. Second, classical measurement error could attenuate the OLS estimates toward zero. Third, reverse causality could also bias the OLS estimates.[1] In consequence, these studies mainly exploit three exogenous sources of variation in parental education: One part of the literature relies on exogenous variation of opportunity costs of schooling (e.g., Currie and Moretti, 2003; Carneiro et al., 2013). A second part of the literature exploits discontinuities created by school entry laws (e.g McCrary and Royer, 2011). The third, and largest, part of the literature exploits changes in the number of years of compulsory education (e.g., Lindeboom et al., 2009; Lundborg et al., 2014; Rawlings, 2015; Chou et al., 2010; Huebener, 2018). While still inconclusive, the majority of these studies point toward a positive effect of maternal education on children's health in infancy and early adolescence. However, the mental health outcomes of children are largely neglected in this literature. To the best of our knowledge, the only exception is Lindeboom et al. (2009) who investigate the effect of changes in the parent's number of compulsory years of education on the mental conditions in early adolescence in the UK. They report no effect of parental education on mental conditions of children in early adolescence. We

[1] Clearly, this would impose a high degree of rationality if the parents adjust their education based on the expectation that this would improve the mental health of their children.

build on the important work of Lindeboom et al. (2009) and extend this work along two important dimensions: first, we focus on children during adulthood. This is important since most mental health symptoms emerge until middle ages (Kessler et al., 2007). Second, we rely on self-reported mental health instead of parental reports of mental illnesses of the children. This is important since parental reports could be biased by either systematic reporting differences between high and low socioeconomic status individuals or the fact that a low socioeconomic status of the parents is potentially associated with undiagnosed health conditions of the children (Case et al., 2002).

In addition, an emerging literature investigates the influence of early life circumstances on adult mental health. Most important for our study, Avendano et al. (2020) estimate the effect of extending compulsory schooling by one year on the students' mental health later in life. Avendano et al. (2020) find that extending years of schooling of those at the lower end of the educational distribution worsens their mental health. They find evidence that this is attributable to the fact that this CSL reform extended schooling for those who did not have a desire for longer schooling. In contrast, analyzing the CSL reform in Germany, Dahmann and Schnitzlein (2019) find no evidence for a protective effect of schooling on one's own mental health.[2] Analogously, Banks and Mazzonna (2012) find a negative albeit insignificant effect of an increase in the compulsory years of education on the respondents' quality of life, assessed by the CASP-19.[3].

More generally, Adhvaryu et al. (2019) investigate the effect of income shocks early in life on mental health in adulthood. In addition, Adhvaryu et al. (2015) show that temperature shocks in utero increase depressive symptoms in adulthood. Almond et al. (2018) provide a thorough review of the complete economic literature regarding the effects of early life influences and in utero experiences on later life outcomes, including mental health.

Our empirical results show that there exists no protective effect of maternal education on children's mental health. However, we can confirm a considerable effect of maternal years of schooling on children's PCS score. The CSL reform increased chil-

[2]In addition, studies examining the relationship between education and life satisfaction are inconclusive.

[3]The CASP-19 consists of 19 items capturing the respondents' control, autonomy, self-realization and pleasure

dren's PCS score by approximately 14% of a standard deviation. This confirms previous studies that find positive effects of the CSL reform on physical health (Huebener, 2018). Similarly, Kemptner and Marcus (2013) find that maternal education affects the physical health of the children. However, they use the number of mothers' siblings as an instrument, identifying a causal effect for a different population.[4] We further investigate the driver of this effect and find that this effect mainly operates through improvements in physical functioning. This is a result that has not been shown in the literature to date. Similar to Huebener (2018), we can also confirm a positive effect on the self-rated health status of the children. However, this effect is smaller in magnitude than the effect on physical functioning.

In the second part of the chapter, we contribute to the literature by identifying potential mediators between maternal education and children's mental health. We do this by performing two separate steps. In a first step, we estimate the effect of the compulsory schooling law reform on a set of potential mediators. In a second step, we check for the partial correlation between these mediators and the MCS score in our main specification. Together, the results provide evidence regarding which mediators can be considered relevant in the relationship of interest. We find that the number of close friends, i.e. social capital, could be a potential mediator of maternal years of schooling on children's mental health. This has not been shown to date. However, the overall contribution of the identified mediators can be considered small, which is consistent with our main result of a zero effect of maternal education on children's mental health.

2.2 Theoretical background

There are a multitude of channels through which maternal education could affect children's health in adulthood. In what follows, we give a short overview regarding which channels we expect that maternal education might affect children's mental health. Since we observe the children mainly in adulthood, this discussion focuses on mechanisms that we are able to test empirically.[5]

[4] In particular, Kemptner and Marcus (2013) compare mothers for whom the parental budget constraint was binding versus those for whom the parental budget constraint was not binding.

[5] The SOEP started the data collection in 1984. Thus, most adult children already left their parental home at this point in time. Therefore, we are not able to test mechanisms on the household

Health behaviors: Empirical studies showed that the correlation between health and socioeconomic status is mainly mediated through health behaviors. For instance, Lynch et al. (1997) find that physical, psychological and cognitive functioning are negatively associated with the number of times individuals find themselves in episodes of economic hardship. They conclude that this relationship is mediated by lifestyle factors such as cooking and physical activities. In an additional example, Chetty et al. (2016) find that differences in life expectancy across commuting zones in the US are predominantly associated with the regional prevalence of (non-)healthy habits such as smoking, obesity and the exercise rate.

Empirical studies show that the gradient in the socioeconomic status and health as well as health behaviors extends into the next generation (Case et al., 2002; Reinhold and Jürges, 2012; Huebener, 2018, 2019). Consequentially, we hypothesize that better educated parents are more likely to raise healthy children and, perhaps even more important, to pass on health knowledge that benefits their children even in adulthood. One example of such health knowledge could be a healthy eating behavior or an active lifestyle, including sport activities. These health behaviors are associated with a body composition that is more healthy and associated with higher confidence.

Empirical studies to date confirm that weight and obesity are negatively associated with mental health (e.g., Simon et al., 2006). However, the evidence for a causal relationship between body composition and individuals' mental health is ambiguous. Willage (2018) investigates the effect of BMI on mental health using an index for genetic risk for a high body mass index (BMI) and the virtue of Mendelian randomization as a source of exogenous variation in the BMI, finding that an increase in the BMI increases suicidal ideation but has no effect on counseling or an index for depression. Similarly, Amin et al. (2020), using the same identification strategy find no effect of BMI on mental health in young adults but in the elderly population. Bargain and Zeidan (2019) find positive and negative effects of the waist-to-height ratio (WHtR) on mental health among the poorest and richest in Mexico, respectively.[6] In conclusion, we conjecture that maternal education could have an effect on the children's mental health via their own body weight.

level when the child was still at home.

[6] According to Bargain and Zeidan (2019), this can be explained by different norms prevailing in different strata of the society.

Education: A person's own education is a central component of a person's socioeconomic status. The health-enhancing effect of education has been proven in a wide range of studies (e.g., Kemptner et al., 2011; Lleras-Muney, 2005). Indeed, the positive effect of education on health is also a focal prediction of the Grossman model for health demand. For instance, education can enhance the efficiency of medical and preventive care as well as of time investments in health (Grossman, 1972), improve the individuals' management of illnesses (Goldman and Smith, 2002), increase the health returns individuals can accrue from new knowledge and facilitate the adoption of new health technologies (Lleras-Muney and Lichtenberg, 2005).

Furthermore, better education is associated with better jobs and higher income. This in turn results in higher job and financial security. Ultimately, improved educational outcomes for children could improve the child's mental health via increased financial and job security. This relationship is also referred to as social causation theory: individuals with better socioeconomic standing enjoy better mental health (e.g., Perry, 1996; Wang et al., 2015).

Equally important, individuals with higher education have been shown to command more psychosocial resources that enhance mental health (e.g., Niemeyer et al., 2019). While theoretical considerations point toward a positive effect of an individual's education on mental health, Piopiunik (2014) shows that the CSL reform caused improvements in the educational outcomes of the children in Germany, predominantly along the mother-daughter line.

However, as alluded to in the introduction, the empirical evidence of a person's education on his or her own mental health is ambiguous with Avendano et al. (2020) and Banks and Mazzonna (2012) pointing toward negative effects of a CSL reform in the UK and Dahmann and Schnitzlein (2019) to zero effects of the CSL reform on individual's mental health in Germany.

Social capital: We hypothesize that maternal education enhances children's social capital and that social capital benefits children's mental health. Social trust, operationalized by the number of close friends, also translates to larger social networks the children could rely on in times of crises. The beneficial effect of social capital on health is well documented (Islam et al., 2006; Petrou and Kupek, 2008; Fujisawa et al., 2009; Hurtado et al., 2011; Fiorillo et al., 2020; Ho, 2016). In addition, education is an important predictor for social capital (e.g., Helliwell and Putnam, 2007; Margaryan et al., 2019). Consequentially, we predict that maternal education improves children's mental health through improved social capital.

Positive assortative mating: Assortative mating is an important empirical phenomenon and is associated with differential availability of household resources and children's outcomes (Greenwood et al., 2014; Eika et al., 2019; Bratsberg et al., 2018; Barban et al., 2019; Holmlund, 2020; Wagner et al., 2020). In general, a partner with higher education can contribute to children's well-being with more financial as well as nonfinancial resources, e.g. health knowledge. Consequentially, we conjecture that maternal education could contribute to children's mental health via the partners' "quality", assessed via partners' educational attainment.

Number of siblings: An alternative channel through which children's mental health could be affected is the number of siblings. Becker and Lewis (1973) contend that parents face a quantity-quality trade-off with respect to their children. Given time and budget restrictions, Becker and Lewis (1973) argue that an increase in the number of children could potentially decrease the "quality" of children, e.g. measured in years of schooling or health. However, the results of empirical studies investigating the quantity-quality trade-off do not unequivocally support or reject it (e.g., Rosenzweig and Wolpin, 1980; Li et al., 2008; Black et al., 2005; Angrist et al., 2005; Briole et al., 2020; Alidou and Verpoorten, 2019; Fernihough, 2017; Mogstad and Wiswall, 2016; Bhalotra and Clarke, 2019). For mental health, studies to date showed either no effect of the number of siblings on children's mental health (Baranowska-Rataj et al., 2016) or a positive correlation (Grinde and Tambs, 2016) between the family size and the children's mental health. In fact, Cygan-Rehm and Maeder (2013) show a negative effect of the CSL reform on maternal fertility, operating mainly through the extensive margin.[7]

[7] Under the conjecture of negative selection into fertility at the extensive margin, our estimate represents a lower bound for a hypothetical positive effect of the CSL reform on children's mental health.

2.3 Method

2.3.1 Institutional background

The German school system. In Germany, children usually enroll in elementary school (*"Grundschule"*) at the age of six. Typically, after four years of elementary school, children continue schooling in one of three different tracks of secondary school. These three tracks differ in terms of duration and curriculum. Among the three school tracks, the basic school track (*"Hauptschule"*) ends after four or five years and provides basic general education. In the intermediate track (*"Realschule"*), children finish school after ten years and experience a more extensive general education. High school (*"Gymnasium"*) lasts nine additional years and offers the most academic curriculum. The basic and intermediate school tracks qualify the student for an apprenticeship or vocational training. In contrast, high schools provide their graduates with the university entrance qualification (*"Abitur"*).[8]

Of note is the allocation to the respective school tracks that depends considerably on the student's academic performance in the main school subjects. The elementary school teacher usually gives a recommendation to the parents, based on the students' aptitude, as reflected in the student's grades. Whereas this recommendation is binding in some states, it is not binding in others. The parents decide which school track their child will attend if the teacher's recommendation is not binding.[9] However, most parents adhere to the teachers' recommendation, even if it is not binding.

Usually, children with the lowest grades are assigned to the basic school track, whereas those with the best grades are assigned to high schools. Mobility between the three school tracks is possible but limited, with downward mobility being more common. Approximately two percent of all students change the track to which they were assigned (Dustmann et al., 2016).

The Compulsory Schooling (CSL) Reform. Starting in the 1940s, the federal states in West Germany implemented CSL reforms that prolonged the number of

[8] In addition, high school also qualifies the student to take up an apprenticeship or vocational training. Over the last decade, reforms changed the length of the high track to eight years and very recently back to nine years for some states. However, the individuals in our sample were not affected by this.

[9] In states with binding teachers' recommendations, parents could also opt for a lower school track.

mandatory schooling years from eight to nine years.[10]

During the Weimar Republic, the educational policy was performed at the state level. This changed after the Nazis were in power. In 1934, the Nazis decreed that the educational policy should be coordinated centrally, on the national level. The goal of this measure was to gain control over the school curriculum (Nicholls, 1978; Tent, 1982; Margaryan et al., 2019). However, this changed after World War 2, when Germany was divided into four occupational zones and the occupying forces aimed at denazifying the school curricula (Tent, 1982). Consequentially, educational policy was performed at the state level thereafter.

During the post-war area, the states decided to prolong the number of mandatory years of schooling from eight to nine years. Broadly, the discourse surrounding the CSL reform can be distinguished into two periods. At the beginning, the main reason to extend the number of compulsory years of schooling was to relieve the strain on the labor market for young individuals. At the beginning of the post-World War 2 era, there existed a shortage of positions as apprentices. As a consequence, the states aimed at holding back the students from the labor market for an additional year. The first state to extend the number of compulsory years of education was Hamburg in 1946. Within the subsequent 16 years, Schleswig-Holstein, Bremen, Lower Saxony and Saarland followed.[11] In the second period, the proponents of the CSL reform advanced the argument that the current students are lacking occupational maturity. For instance, the average job profile became more demanding and employers increased their expectations toward the basic skill set of their prospective apprentices, e.g. reading, writing and math. Consequently, the states jointly formalized the CSL reform in the *Hamburg Accord*. Eventually, North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Baden-Wuerttemberg implemented the reform in 1967. Bavaria followed in 1969 (see Pischke and von Wachter (2008) for more details).[12] This allows us to exploit

[10]This comprises the Federal Republic of Germany without the states of the former German Democratic Republic and Berlin. A complete list of states that performed the CSL reform is provided in Table 2.1.

[11]Margaryan et al. (2019) test whether the pre-reform unemployment rate and GDP per capita can explain the reform timing and find that these characteristics are not significantly related to the timing of the CSL-reform.

[12]The description of the reform details follows Cygan-Rehm and Maeder (2013) and the reference to LeSchinsky and Roeder (1980) therein. The reform details differ from the ones depicted in Pischke and von Wachter (2008), who rely on LeSchinsky (1981) and Petzold (1981). The details of the CSL reform between Cygan-Rehm and Maeder (2013) and Pischke and von Wachter (2008) differ for the smallest states, which are Hamburg, Schleswig-Holstein, Bremen and Saarland. In Cygan-Rehm

Table 2.1: Details of the timing of the CSL reform in Western German states

	First year all students were supposed to graduate after a minimum of nine school years	First birth cohort affected by change in compulsory schooling law
Hamburg	1946	1931
Schleswig-Holstein	1947	1932
Bremen	1959	1944
Lower Saxony	1962	1947
Saarland	1958	1943
North Rhine-Westphalia	1967	1953
Hesse	1967	1953
Rhineland-Palatinate	1967	1953
Baden-Wuerttemberg	1967	1953
Bavaria	1969	1955

Note: Details stem from Cygan-Rehm and Maeder (2013) and LeSchinsky and Roeder (1980).

exogenous variation across time and space.

2.3.2 Empirical strategy

In our empirical analysis, we are interested in the following relationship:

$$\begin{aligned} MH_i = \beta_1 + \beta_2 YOS_i + \beta_3 age_i + \beta_4 age_i^2 + +\beta_5 \mathbb{1}[i \text{ is female}] \\ + \sum_{t=2002}^{2012} \lambda_t \mathbb{1}[t = t_i \cap t \text{ is even}] + \sum_{s=1}^{9} \mu_s \mathbb{1}[s = s_i] \\ + \sum_{s=1}^{9} \rho_s \mathbb{1}[s = s_i] c_i + \sum_{c=1930}^{1960} \kappa_c \mathbb{1}[c = c_i] + \epsilon_i, \end{aligned} \tag{2.1}$$

according to which the mental health outcome MH of child i is a function of the maternal years of schooling, YOS_i. In addition, we control for a second-order polynomial in age of the children to adjust for the age-mental health profile. Further, to adjust permanent gender differences in mental health, we include an indicator for being female, with $\mathbb{1}[\cdot]$ being the notion of an indicator function. Last, we also control for contemporaneous shocks in the children's mental health on the year level t, λ_t, by including survey year indicators.

As previously alluded to, the OLS estimate of the coefficient of interest, β_2, is most likely inconsistent. Two factors could render our estimate inconsistent: First, unobserved factors could bias the relationship of interest. For example, mothers of higher ability, on average, accrue more years of schooling and raise healthier children.

(2018), the author compares the reform details in both sources with official educational statistics and concludes that LeSchinsky and Roeder (1980) corresponds most closely to the official statistics.

In consequence, we would expect the OLS estimate of β_2 to be biased upward. Second, classical measurement error in the maternal years of schooling, YOS_i, could attenuate the magnitude of our estimates toward zero. As a result of this classical measurement error, we would underestimate the strength of the relationship of interest. To account for the potential endogeneity of the maternal years of schooling, we instrument it with an indicator that is equal to one if the mother has been exposed to the CSL reform, $\mathbb{1}$[mother of i exposed to CSL reform], as displayed in Equation 2.2:

$$\begin{aligned} YOS_i = \gamma_1 + \gamma_2 \mathbb{1}[\text{mother of } i \text{ exposed to CSL reform}] + \gamma_3 age_i + \gamma_4 age_i^2 \\ + \gamma_5 \mathbb{1}[i \text{ is female}] + \sum_{t=2002}^{2012} \omega_t [t = t_i \cap t \text{ is even}] \\ + \sum_{s=1}^{9} \delta_s [s = s_i] + \sum_{s=1}^{9} \zeta_s [s = s_i] c_i + \sum_{c=1930}^{1960} \pi_c [c = c_i] + \eta_i. \end{aligned} \tag{2.2}$$

The excluded instrument in our 2SLS framework is $\mathbb{1}$[mother of i exposed to reform]. As additional included instruments, we include maternal year of birth, state of schooling and linear state of schooling-specific trends in the maternal year of birth in Equations 2.1 and 2.2. Thus, we exploit variation within the maternal state of schooling and year of birth as identifying variation rather than permanent differences across states and cohorts. Throughout, robust standard errors are clustered on the level of the policy assignment (Bertrand et al., 2004), which in our case is the maternal state of schooling.[13]

As we will explain in more detail in our data section, our data allows us to consistently distinguish between mental and physical health. This allows us to contrast the results on mental health by also testing for effects on the physical health dimension.

[13]We also calculate p-values based on the wild cluster bootstrap-t procedure. The reason is that we only have ten clusters and the asymptotic results for clustered variance-covariance matrices rely on the number of clusters converging to infinity. As a consequence, standard errors, clustered on the maternal state of schooling, of which we have only ten, are most likely inconsistent (Cameron et al., 2008; Cameron and Miller, 2015). Two-way clustering, e.g. on the maternal state of schooling and year of birth-level, is no solution to this problem. Since the asymptotic results for clustered variance-covariance matrices in case of two-way clustering depend on the number of clusters of the clustering dimension with the fewest number of clusters, two-way clustering would still result in inconsistent standard errors in our case (Cameron and Miller, 2015). Statistics based on the wild cluster bootstrap-t procedure converge faster because of their asymptotic refinement (Cameron et al., 2008). Therefore, we will provide consistence inference for all our results in the robustness section and thereafter. Throughout, we will use the **boottest** command in Stata 15 (Roodman et al., 2019).

For this, we will also apply our 2SLS framework to the physical health dimension, as captured in the PCS score.

After estimating the effect of the additional year of maternal schooling on the child's mental health, we infer the channels through which this additional year of maternal schooling affects their child's mental health. We restrict our attention to mediators on the child's level as well as parents' time-invariant characteristics since we are not able to observe time-variant characteristics while the children are in their parental household.[14] As discussed above, our mediators on the child's level are proxies for

- body composition (BMI and likelihood of obesity),
- human capital (educational attainment),
- and social capital (number of friends).

In addition, our proxies for changes in the time-invariant home environment are the following:

- assortative mating (father's educational attainment),
- human capital of the mother (mother's vocational education),
- and maternal fertility (number of children).

We check for these mediating channels by applying our 2SLS framework on the potential mediators.[15] In a second step, we include our potential mediators in Equation 2.1 to check for the partial correlation of the mediators with the mental health outcome. If both coefficients are significant and the explanatory variable changes size, we have evidence for a potential mediator (Baron and Kenny, 1986). The product of these two coefficients informs us about the direction and the magnitude of the overall contribution to the relation of interest.

[14]The first wave of the SOEP was administered to individuals in 1984.

[15]We do not include the children's age and the survey year indicators if we test for the maternal outcomes. In addition, we only take the first observation if we check for the maternal mediators which is relevant because mothers often have one or more children.

2.3.3 Identification

For the 2SLS estimate of β_2 to be consistent, we need the following assumptions to hold: *instrument independence*, *exclusion restriction* and *instrument relevance*. Further, we have to be able to adhere to the *stable unit treatment assignment (SUTVA)*, and under the plausible assumption of effect heterogeneity, we require the nonexistence of defiers. As consequence of the latter *monotonicity assumption*, we will estimate the *local average treatment effect (LATE)*. This is the effect for individuals who comply with the treatment. In our setting, the LATE coincides with the effect for individuals whose mothers' years of schooling increased due to the CSL reform but otherwise would have not in the absence of the CSL reform. These compliers are students from the basic school track since those that are always takers in the other two school tracks already attend more than eight years of schooling.

Independence assumption: The independence assumption is met if the instrument is independent of the vector of potential outcomes and potential assignment to an additional year of schooling for the mothers. To put it differently, the maternal reform status and the children's mental health should not share a common cause. Thus, the validity of the instrument would be threatened if reasons or the timing of the CSL reform is associated with the children's mental health.

We argue that the independence assumption is met for three reasons: First, the main actor in German health policy is the federal government. The role of the states is limited to implementing the laws and ordinances of the federal government. Other areas comprise the supply of hospitals and preventive care measures.

The second reason why we argue the independence assumption is met is the fact that the curriculum of the additional school year, due to the CSL reform, typically does not comprise any content to improve the students health. As Petzold (1981) contends, the CSL reform's goal was to improve the students' occupational maturity. For example, the reasons to introduce the 9th school year in North Rhine-Westphalia, the largest federal state in Germany, were (i) the provision of a better general education, (ii) deepening the political education, and (iii) the acquisition, practice and expansion of basic skills and knowledge (Margaryan et al., 2019).

Third, to account for potential differential trends in state-level characteristics that might be associated with the timing of the implementation of the CSL reform and the children's mental health, we include linear state-specific trends in the maternal year of birth in Equations 2.1 and 2.2.[16]

Exclusion restriction: The exclusion restriction requires that the CSL reform altered children's mental health solely through the maternal years of schooling. One way through which the exclusion restriction would be violated is if the CSL reform would have altered the health care provision, for instance. However, again, health policy is mainly performed on the level of the federal government.

An alternative channel trough which the exclusion restriction could be violated by the CSL reform is changes in the tracking regime (Lundborg et al., 2014). For example, a postponement of tracking in school could have changed the pool of potential partners. In Germany, school tracking typically happens after the fourth grade and was not changed by the CSL reform. Thus, we are not concerned with violations of the exclusion restriction because of changes in school tracking.

Similarly, the CSL reform could have been associated with a decline in teaching quality. For instance, if the implementation of the CSL reform was not accompanied by an increase in the number of teachers, the teaching quality could have been altered, e.g. by an increase in class sizes. We address this issue in the robustness section and do not find that our results are altered by these organizational challenges associated with the CSL reform. Thus, we conclude that the exclusion restriction is met in case of the CSL reform in Germany.

Instrument relevance: The instrument relevance requires that the instrument is significantly associated with the endogenous regressor. If this were not the case, even minor violations of the exclusion restriction could introduce major biases (Bound et al., 1995). In contrast to most of the other assumptions, the relevance of the instrument is testable. Typically, a test of the significance of the excluded instrument in the first stage should yield an F-statistic of ten or larger (Staiger and Stock, 1997). This requirement is met in all our regressions.

Monotonicity: Under the plausible assumption of effect heterogeneity, we have to be able to rule out defiers (Angrist et al., 1996). Defiers are students who would continue schooling in absence of the CSL reform but would stop schooling after eight years after the CSL reform was implemented. Since eight years of schooling was mandatory

[16]In the robustness section, we also include a second-order polynomial in the maternal state of schooling trend to allow for more flexible associations.

before the implication of the CSL reform and nine years after the implementation of the CSL reform, we are confident that defiers can be ruled out.

SUTVA: The SUTVA requires that the effect of a treatment depends only on the individual's treatment assignment and not the assignment of any other individual. This basically rules out equilibrium effects. For instance, if the number of years of schooling would convey meaningful information about the student's latent ability, this signaling value would vanish if every student were to take an additional year of schooling at the relevant margin. This could in turn change the effect of mothers' schooling on children's mental health. However, this is a limitation every study exploiting schooling reforms shares and which is difficult to address without explicitly modeling the relations of interest and additional assumptions.

2.4 Data

The data uniquely suited for our task are from the SOEP. The SOEP is a representative longitudinal study of private households in Germany. It started in 1984 and surveys – in the most recent wave – approximately 15,000 households and more than 25,000 persons living in Germany annually.[17] Among other topics, the panel covers household composition, occupational biographies, employment, earnings, health, and parenting behaviors (Göbel et al., 2018).

Moreover, children in the household that are surveyed in the SOEP are surveyed first in the year they become 17 years old and are interviewed annually thereafter, even after leaving the parental household. Hence, the SOEP entails detailed information about mother-child pairs. Consequently, we are able to link the educational information of SOEP respondents to their children's mental health outcomes.

Starting in 2002, the SOEP introduced a special health module that is administered biannually since then. This survey module includes the SF-12 questionnaire, which comprises twelve health-related questions that cover both mental and physical health dimensions. The items are displayed in Figure 2.B.1 in the appendix. The questions refer to the health status within the four weeks preceding the interview. Hence, the SF-12 questionnaire refers to the current health status (Andersen et al., 2007). The

[17]We use SOEPv32. DOI: 10.5684/soep.v32

MCS score is the second factor of a principal component analysis of the SF-12 questionnaire in the 2004 SOEP population and is reflective of the individual's mental health (Andersen et al., 2007). Typically, the MCS score is normalized to have mean 50 and standard deviation 10 in the 2004 SOEP population.[18] Higher MCS scores indicate better mental health. The MCS score is widely used in the epidemiological literature and has high predictive power for mental illnesses (Salyers et al., 2000). The MCS score also has been widely used in the economic literature in recent years (e.g., Marcus, 2013; Eibich, 2015; Cygan-Rehm et al., 2017).

Along with the MCS score, we also use an indicator for being at risk of developing symptoms of a mental disorder based on the MCS score. An MCS score below 45.6 has high predictive power for the occurrence of symptoms of a clinically relevant mental disorder over a range of thirty days (Vilagut et al., 2013).[19] Based on this threshold, we derive our indicator for being at risk of a mental disorder.

A distinct advantage of the SF-12 questionnaire is that it infers the mental health status indirectly. Direct measures, such as a self-reported diagnosis of mental illnesses, are prone to underreporting. For example, Bharadwaj et al. (2017) show that individuals tend to underreport mental illnesses in 36% of the cases in surveys compared to administrative data on diagnoses. Since the SF-12 questionnaire is not asking for mental health or illnesses directly in any item, we are confident that underreporting is not a concern in our context.

One further advantage of the SF-12 questionnaire is that it also allows us to contrast the effect of the maternal years of schooling on the children's mental health with the effect on children's physical health in a consistent way. For this, we use the PCS score as it stems from the same principal component analysis as the MCS score. To our knowledge, such a systematic and internally consistent comparison of the effect of maternal schooling on the children's mental and physical health has not taken place in the literature to date.

To construct the CSL reform indicator, we need information on the place of schooling and the year of birth of the mother.[20] Fortunately, the SOEP entails information

[18]The MCS and PCS scores are calculated by the SOEP group and are part of the data provision.

[19]A specificity of 86% is associated with the threshold of 45.6. That is, 86% of the individuals with an MCS score below 45.6 actually exhibit symptoms of a clinical relevant mental disorder within 30 days preceding the screening (Vilagut et al., 2013).

[20]We assume that all children enter school in the year they become six years old. Since we do not observe the level of compliance with the school enrollment guidelines among our sample, as a

about the last state of schooling. This information is available for approximately 20% of the SOEP sample. For the remaining mothers, we take the state in which the mothers were living when they were first surveyed.[21] This procedure is consistent with the procedure applied in previous studies exploiting the CSL reform in Germany (e.g., Pischke and von Wachter, 2008; Cygan-Rehm and Maeder, 2013).

Similar to most datasets in Germany, the SOEP does not contain a measure for the number of years of schooling. However, we observe the school-leaving degree of the individuals. Therefore, we impute the number of years of schooling. We assign individuals that have a basic or no school-leaving degree of eight years of schooling if they are not exposed to the CSL reform and of nine years if they are exposed to the CSL reform. We assign individuals with an intermediate school-leaving degree of ten years of schooling. Those with an academic school-leaving degree are assigned 12 or 13 years of schooling, depending on whether they obtained the "*Fachhochschulreife*" or "*Abitur*". This procedure is established in the literature; compare, for instance, Pischke and von Wachter (2008) and Cygan-Rehm and Maeder (2013).

We restrict our sample to the mother-child pairs for which we have at least one observation on the children's mental health outcomes from 2002 through 2014. If we have multiple observations on the mental health outcomes, we choose the observation that is closest to the age of 35. The reason for this age choice is that this corresponds approximately to the median age at which common mental disorders emerge (Kessler et al., 2007). In addition, we keep information for mother-child pairs for which we have information about the mother's years and place of schooling in addition to the mother's year of birth. We exclude observations of mothers who migrated to Germany after 1945.[22] Further, we restrict our sample such that our mothers are born in the year range 1930-1960, comparable to the procedure in Cygan-Rehm and Maeder (2013) and Margaryan et al. (2019). The summary statistics for our sample are reported in Table 2.2.

robustness test, we drop the first cohort that is affected by the CSL reform. Our estimates are robust to this test.

[21]For 83.46% of the respondents in the SOEP, the stated state of schooling corresponds to the first state in which the respondents, who stated their state of schooling, were first surveyed. If we condition on individuals with a school-leaving degree from the basic track, our complier group, this number increases to 90.76%.

[22]As a result, our sample of mothers consists of mothers who have no migration background and mothers who have parents that were not born in Germany but were born in Germany themselves.

Table 2.2: Summary statistics

	Not affected by CSL reform		Affected by CSL reform	
	(1) Mean	(2) N	(3) Mean	(4) N
Panel A: Children				
MCS Score	49.031	1458	49.482	1873
	(9.397)		(9.697)	
Being at risk of mental disorder	0.310	1458	0.295	1873
PCS Score	54.524	1458	55.649	1873
	(7.361)		(6.838)	
SF-12 subscales				
Mental health	49.384	1458	49.595	1873
	(9.215)		(9.645)	
Role emotional	50.923	1458	51.849	1873
	(9.590)		(9.294)	
Social functioning	50.914	1458	51.496	1873
	(9.521)		(9.266)	
Vitality	51.198	1458	51.927	1873
	(9.344)		(9.648)	
General health	53.450	1458	54.673	1873
	(8.795)		(8.803)	
Bodily pain	52.444	1458	53.333	1873
	(8.727)		(8.278)	
Role physical	52.624	1458	53.981	1873
	(8.787)		(8.138)	
Physical functioning	51.198	1458	51.927	1873
	(9.344)		(9.648)	
Female	0.440	1458	0.482	1873
Age	32.594	1458	24.754	1873
	(6.128)		(5.238)	
Year of birth	1973.464	1458	1984.819	1873
	(8.194)		(5.780)	
School-leaving degree				
Basic	0.241	1390	0.178	1507
Intermediate	0.297	1390	0.305	1507
High	0.462	1390	0.517	1507
BMI	24.687	1453	23.702	
	(4.235)		(4.313)	
Obese	0.099	1453	0.081	1861
Number of friends	4.859	834	5.016	1453
	(3.450)		(3.431)	
Panel B: Mother				
Years of schooling	9.134	1458	10.434	1873
	(1.641)		(1.566)	
School-leaving degree				
Basic	0.609	1458	0.360	1873
Intermediate	0.265	1458	0.358	1873
High	0.126	1458	0.282	1873
Vocational degree	0.715	1448	0.885	1868
Number of children	2.567	1440	2.382	1718
	(1.237)		(1.177)	
Year of birth	1944.157	1458	1956.110	1873
	(6.603)		(3.489)	
Partner's school-leaving degree				
Basic	0.562	1259	0.374	1559
Intermediate	0.179	1259	0.217	1559
High	0.260	1259	0.409	1559

Note: Table 2.2 displays summary statistics for children and their mothers, conditional on maternal reform exposure. The PCS and MCS score correspond to the first and second factor of a principal component analysis of the items of the SF-12 questionnaire. The indicator for being at risk of a mental disorder is equal to one if the respondents MCS score is below 45.6. Numbers in parentheses are standard deviations for continous variables.

2.5 Results

2.5.1 Maternal schooling and children's mental health

The main results are displayed in Table 2.3. Panel A displays the results for the MCS score, Panel B presents results for the indicator for being at risk of having a mental disorder and Panel C for the PCS score. Columns (1) and (2) of Table 2.3 display OLS regressions based on Equation 2.1, with the difference that the OLS regression of column (1) contains no age and survey year controls whereas column (2) does. Columns (3) and (4) replicate the specifications in columns (1) and (2), respectively, in our 2SLS framework. The corresponding first-stage results are displayed in the bottom Panel.[23]

Our OLS specifications in columns (1) and (2) indicate that there exists no association between the maternal years of schooling and the mental health of the children. In our preferred specification, displayed in column (2), the coefficients are small and insignificant. Moreover, the result in column (2) of Panel B implies virtually no association with the indicator for being at risk of having a mental disorder. In contrast, and in line with the abovementioned literature, the maternal years of schooling are positively associated with the children's PCS score. One year of maternal years of schooling is associated with a 5% standard deviation increase in the child's PCS score.

In a next step, we apply our 2SLS framework. The first-stage results are displayed in the bottom panel of Table 2.3. The F-statistic in our preferred specification is 25.68 and suggests the absence of a weak instrument (Staiger and Stock, 1997; Stock and Yogo, 2005).

Column (4) of Table 2.3 displays the 2SLS estimate of the coefficients on the maternal years of schooling for our preferred specification. Panel A suggests that there exists no effect of maternal years of schooling on the MCS score. Compared to the OLS estimate, displayed in column (2), the coefficient estimate is almost identical and not significant. Further, we also find no effect on our indicator for being at risk of a mental disorder.

Conversely, we find a positive effect on the PCS score, as displayed in our main

[23]Throughout, we display Kleibergen-Paap rk Wald F-statistics. However, we abbreviate it as F-statistic.

Table 2.3: Maternal years of schooling and children's mental and physical health

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Panel A: MCS score				
Maternal years of schooling	-0.001	-0.014	-0.004	-0.015
	(0.012)	(0.013)	(0.113)	(0.105)
Panel B: Being at risk of having a mental disorder				
Maternal years of schooling	-0.005	0.000	-0.053	-0.049
	(0.004)	(0.004)	(0.067)	(0.065)
Panel C: PCS score				
Maternal years of schooling	0.060***	0.050***	0.164***	0.140**
	(0.006)	(0.006)	(0.060)	(0.060)
First stage				
Maternal reform exposure			0.789***	0.777***
			(0.120)	(0.153)
F-statistic			43.361	25.679
Observations	3331	3331	3331	3331
Age controls and survey year FE	✗	✓	✗	✓

Note: Table 2.3 displays OLS and IV estimates for a regression of the mental health outcomes of the children on maternal years of schooling. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth, a state of schooling specific linear trend in the maternal year of birth and an indicator which is equal to one if the child is female. Robust standard errors, in parentheses, are clustered on the maternal state of schooling level. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

specification in column (4) of Panel C. According to that estimate, the CSL reform increased the children's PCS score by approximately 14% of a standard deviation. Comparing this effect size to previous studies, we acknowledge that comparisons are complicated by the fact that previous papers focused on proxies for health, such as health behavior, e.g., smoking and body composition (BMI), or diagnoses. For Germany, Huebener (2018) finds that CSL reform decreases the likelihood that the children smoke currently by approximately 26%, decreases the BMI by approximately 3.4% and the likelihood of having a chronic condition by approximately 22%. Looking at self-rated health, Huebener (2018) concludes that the reform increased the self-rated health status by approximately 9.6% of a standard deviation. Lundborg et al. (2014) find that a CSL reform in Sweden increased children's global health and height by approximately ten percent of a standard deviation. Thus, our effect size is lower than comparable effects on health behaviors and is in a similar range to the effect on self-rated health status and height, with the latter being an established proxy for global health.

A further investigation of the subscales of the SF-12 questionnaire further supports the conclusion that there indeed exists no effect on the mental health of children. The results are presented in Table 2.4. Panel A depicts the results for the 2SLS regression of the subscales that are strongly associated with the MCS score. Panel B shows the comparable results for the subscales that are strongly associated with the PCS score. The first-stage results mirror the ones in Table 2.3. As Panel A shows, there is no significant effect of the CSL reform on any of the subscales associated with the MCS score. In contrast, we find a significant effect on the subscales "Physical functioning" and "General health". The CSL reform increased physical functioning by approximately 17.5% of a standard deviation. The CSL reform increased "General health", which corresponds to the self-rated health status, by approximately 5.3% of a standard deviation. This is consistent with the finding of Huebener (2018), who also finds an effect of maternal years of schooling on self-rated health.

Table 2.4: Maternal years of schooling and the children's SF-12 subscales

	(1)	(2)	(3)	(4)
Panel A: MCS score	Mental health	Role emotional	Social funct.	Vitality
Maternal years of schooling	-0.070	0.054	0.097	0.019
	(0.076)	(0.093)	(0.126)	(0.109)
Panel B: PCS score	General health	Bodily pain	Role physical	Physical funct.
Maternal years of schooling	0.053*	0.077	0.049	0.175***
	(0.031)	(0.067)	(0.066)	(0.058)

Note: Tables 2.4 displays the results of a 2SLS regression of the SF-12 subscales on maternal years of schooling. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth, a state of schooling specific linear trend in the maternal year of birth, a second order polynomial in the children's age, survey year indicators and an indicator which is equal to one if the child is female. Panel A displays results for the subscales which are strongly associated with the MCS score. Panel B displays results for the subscales which are strongly associated with the PCS score. Robust standard errors, in parentheses, are clustered on the maternal state of schooling level. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

2.5.2 Robustness tests

Inconsistent inference: One concern could be, that our variance-covariance matrix is inconsistent because of the small number of clusters, of which we have ten (Cameron et al., 2008; Cameron and Miller, 2015). Since asymptotic theory for clustered variance-covariance matrices relies on the number of clusters converging to infinity, we conclude that our standard errors might be inconsistent.[24] Therefore, we apply wild cluster

[24]Two-way clustering is not a solution to this problem since the effective number of clusters coincides with the number of clusters of the clustering dimension with the smallest number of clusters (Cameron and Miller, 2015).

bootstrap procedures with Rademacher weights, 999 repetitions and equal tail p-values to account for potential violations of the symmetry assumption of the distribution of the t-values, which could be the case because IV estimates tend to be biased toward the corresponding OLS estimate (Cameron and Miller, 2015; Roodman et al., 2019). Panel A of Table 2.5 shows the results. It replicates the 2SLS estimates of our main results and the corresponding clustered standard errors. In squared brackets, we display the p-values stemming from the wild cluster procedure. As it can be seen, the level of significance remains unaltered accounting for potential inconsistencies. In what follows, we will always display p-values of the wild cluster bootstrap-procedure along with our results.

Organizational problems: A possible threat to our identification strategy is the fact that the states that were the first to implement the CSL reform might have faced organizational problems. Examples of such organizational problems include a shortage of teachers during that period and a decreased teacher-to-student ratio. These organizational problems could have affected the school quality, which would invalidate the exclusion restriction of our 2SLS strategy. One potential remedy for this threat is dropping those states first implementing the CSL reform. These are Hamburg and Schleswig-Holstein. Other states followed with a twelve-year difference. Based on the conjecture that this time span is long enough to prepare for the CSL reform, we drop all observations from Hamburg and Schleswig-Holstein and perform the main regressions in our favorite specification. The results are presented in Panel B of Table 2.5. Similar to our main results, we observe that maternal years of schooling have no effect on the children's mental health but a positive effect on the children's physical health.[25]

Measurement error in the assignment to the CSL reform: Students could have been wrongly assigned to the CSL reform. This could have happened if the enrollment into school relies not only on the year but also on the month of birth. Likewise, states and individuals could have complied only imperfectly with the CSL reform for the first affected birth cohorts. In addition, the implementation of the

[25]For the wild cluster bootstrap procedure, we rely on weights from the Webb six-point distribution. The reason is that if the number of clusters is equal to eight, there exists 2^8 possible bootstrap samples. This is smaller than 999. Thus, the expected share of draws that mirror the original sample is 2^{-8}. It is unclear how to calculate a bootstrap p-value in this situation. Relying on Webb's six-point distribution minimizes this risk (Roodman et al., 2019).

CSL reform could have deteriorated the schooling quality in particular for the first-affected cohorts in the respective states. Hence, we check the degree to which these circumstances violated our research design, we drop the first cohort affected in each state. The results are displayed in Panel C of Table 2.5. Our main results are not affected: The effect of maternal schooling on mental health is zero. However, for the PCS score, the effect size decreases by approximately 27.7% but remains significant on a ten percent level of significance, relying on conventional inference. However, if we consider the p-value from the wild cluster bootstrap procedure, the effect is close to significance with a p-value of 0.104.

World War 2 Trauma: The exposure to World War 2, which ended in 1945, could have traumatized mothers. Thus, we drop mothers born before 1946. It is worth mentioning that we lose most of the cross-sectional variation in the instrument and that, as a consequence, our identification mostly relies on variation in time in this setting. This is because during the first period of the CSL reform, the states idiosyncratically decided to implement the CSL reform at different points in time and did not coordinate like the states in the second period of the CSL reform. Now, we observe that all estimates increase in magnitude. Notably, this change in effect sizes is most likely reflective of the lack of cross-sectional variation in the instrumental variable, i.e. the estimates are confounded by a common time trend in schooling and mental health..[26] In light of the loss of variation in the instrumental variable, the wild cluster bootstrap p-value is 0.132 now for the MCS score, indicating that the effect is not significant.

Violation of the independence assumption: If the timing of the CSL reform and the mental health of the children share a common cause, the independence assumption would not be met. Margaryan et al. (2019) test whether the timing of the reform is correlated with various (pre-reform) socioeconomic characteristics as well as characteristics for school quality on the state level. They find no evidence that any of these characteristics is associated with the timing of the CSL reform. To account further for differential developments across states in school quality, GDP or other factors potentially determining the timing of the CSL reform and the children's

[26] In a robustness test, Kemptner et al. (2011) restrict the selection of states to those that implement the CSL reform jointly in the second period of the CSL reform to minimize concerns that mobility patterns between states threaten identification. They also observe that point estimates increase in magnitude.

health, we already include state-specific trends in the maternal year of birth. To allow for more flexibility in the functional form of these trends, we also allow for squared state-specific trends. The results are presented in Panel E of Table 2.5. While the coefficients increase in magnitude, the conclusions again remain unaltered.

Table 2.5: Robustness checks

	(1) MCS	(2) Risk ment. disorder	(3) PCS	(4) F-stat.	(5) Observations
Panel A: Consistent inference					
Maternal years of schooling	-0.015	-0.049	0.140**		
	(0.105)	(0.065)	(0.060)	25.679	3331
	[0.915]	[0.813]	[0.034]		
Panel B: Dropping the first two affected states (Hamburg and Schleswig-Holstein)					
Maternal years of schooling	-0.032	-0.035	0.136**		
	(0.106)	(0.062)	(0.056)	33.592	3126
	[0.881]	[0.931]	[0.046]		
Panel C: Dropping the first cohort affected in each state					
Maternal years of schooling	-0.036	-0.045	0.099*		
	(0.137)	(0.088)	(0.058)	14.914	3170
	[0.843]	[0.847]	[0.104]		
Panel D: Dropping cohorts born before 1946					
Maternal years of schooling	-0.122**	-0.003	0.245**		
	(0.062)	(0.043)	(0.104)	18.053	2546
	[0.132]	[0.975]	[0.064]		
Panel E: Squared trend in maternal year of birth for each maternal state of schooling					
Maternal years of schooling	-0.062	-0.020	0.206***		
	(0.041)	(0.037)	(0.072)	27.283	3331
	[0.240]	[0.701]	[0.060]		

Note: Table 2.5 displays IV estimates for a regression of the mental health outcomes of the children on maternal years of schooling, instrumented by maternal reform exposure. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth, a state of schooling specific linear trend in the maternal year of birth, a second order polynomial in the children's age, survey year indicators and an indicator which is equal to one if the child is female. Wild-cluster bootstrap p-values, based on 999 repetitions and Rademacher weights, for the corresponding estimates are in squared brackets. For Panel B, we use Webb weights. Robust standard errors, in parentheses, are clustered on the maternal state of schooling level. Significance stars are based on p-values based on the non-bootstrapped standard errors. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

2.5.3 Potential mediators

In the previous subsection, we confirmed the absence of an effect of maternal years of schooling on the children's mental health. However, the absence of any effect does not mean that their exist no mediators between maternal years of schooling and children's mental health. For instance, two mediators canceling each other off would be consistent with an overall null effect. In what follows, we test which factors could or could not mediate the relationship of interest. For that, we first estimate the effect of maternal years of schooling on outcomes on the children's or maternal level. Then, we test for the partial correlation of the potential mediators in our main 2SLS specification. In combination, these two estimates can inform us about those mediators in our relation of interest.[27]

Weight and obesity

We follow the World Health Organization and define a person as obese if the person's BMI is above 30 (WHO, 2020). Our results indicate neither an effect of maternal schooling on the child's BMI nor on the likelihood that the child is obese, as depicted in Panel A of Table 2.A.1. The effect sizes are small and not significant throughout. Thus, we can rule out children's body weight as a mediator between maternal schooling and mental health.

Child's own educational attainment

In Panel B of Table 2.A.1, we estimate the effect of maternal schooling on the child's educational attainment. The results indicate the absence of an effect of maternal years of schooling on the child's education.[28]

In addition, we find no association with the child's own educational attainment with the child's mental health outcome in our 2SLS framework, as presented in Panel B of Table 2.A.3. In conclusion, together with the finding of Dahmann and Schnitzlein (2019), who find no effect of one's own education on mental health, we can rule out the child's own educational attainment as a mediator between maternal years of schooling and the child's mental health.

[27] Unfortunately, we do not have information on maternal outcomes during the children's childhood. The SOEP started in 1984. In that year, most children in our sample already left their parental home. However, instead, we focus on time-invariant outcomes on the maternal level. Throughout, we present conventional clustered standard error and wild cluster bootstrap p-values.

[28] However, we find that one additional year of maternal schooling increases the educational outcomes of the sons. For instance, the likelihood that that the sons have an intermediate school-leaving degree increases by approximately seven percentage points. This result is consistent with Piopiunik (2014) who also finds that maternal years of schooling has a positive effect on sons' educational attainment but not on the daughters' educational attainment. The results are available upon request.

Social capital

We find a positive effect of maternal years of schooling on the child's number of friends, shown in Panel C of Table 2.A.1.

In Table 2.A.3, we test for the partial correlation of the number of friends with children's mental health in our main 2SLS specification. The association of the children's mental health with the number of close friends is indeed positive and significant. However, the implied overall effect of one additional close friend corresponds to less than 1% ($= 0.012 * 0.733 * 100\%$) of a standard deviation. Thus, while the number of close friends is a mediator in the relationship of interest, the implied effect is rather small. Thus, our finding is consistent with a zero effect of maternal years of schooling on the child's mental health.

Positive assortative mating

We investigate patterns of assortative mating by estimating the effect of maternal schooling on the partners' educational attainment, depicted in Table 2.A.2. We find indeed a positive effect on the partners' education.

However, as in Table 2.A.3, neither the partners' years of schooling nor the fact that the partner has a school-leaving degree higher than the basic school-leaving degree exerts an independent association with the children's MCS score. Thus, we can rule out the fathers educational attainment as a mediator.

Mothers' vocational degree

The CSL reform did not result in a different school-leaving degree for the mothers. However, it could have potentially altered the likelihood that the mothers obtain a vocational degree.

The German labor market is highly formalized. Since the reform of the compulsory years of schooling is not associated with a higher formal school-leaving certificate, it is very unlikely that the career path of the mother is altered in response to the CSL reform only. However, an alternative channel could be that mothers who were exposed to the CSL reform obtained a vocational degree but would not have obtained a vocational degree in absence of the CSL reform. A vocational degree could be associated with higher incomes, among others.

However, we do not find an effect of the CSL reform on the likelihood that mothers have a vocational degree, as depicted in Table 2.A.2.[29] In addition, Table 2.A.3 suggests that the vocational degree of the mother is not associated with the children's MCS score.

Number of children

One way through which maternal years of schooling can affect the child's (mental) health is through the number of siblings. Table 2.A.2 shows that the CSL reform does not affect the number of children in the household. This stands in contrast to the fertility effects found by Cygan-Rehm and Maeder (2013). One reason for the difference could be the fact that we condition on a sample of mothers, while Cygan-Rehm and Maeder (2013) considers all women, regardless of whether they have children or not. Thus, the effect of the CSL reform in Cygan-Rehm and Maeder (2013) is a combination of fertility at the intensive and extensive margin. The combined effect in Cygan-Rehm and Maeder (2013) is approximately 6%. However, they find that the effect at the extensive margin is 20%. Since the overall effect in Cygan-Rehm and Maeder (2013) is a combination of the effect on the intensive and extensive margin, our results appear consistent with Cygan-Rehm and Maeder (2013). In addition, Huebener (2018), similar to our study, conditions on a sample of mothers and finds only a small negative effect on the number of children of mothers. The implied effect size is not larger than 2.2 % and significant at the ten percent level of significance. Further, the implied 95% confidence interval of our estimate does include the estimated effect in Huebener (2018), which is approximately -0.05.[30] We conclude that the fertility effect mainly operates at the extensive margin.

In addition, consistent with the previous epidemiological literature, we find that the number of children in a family is positively associated with the child's MCS score, as shown in Panel F in Table 2.A.3. However, since we are not able to distinguish our 2SLS coefficient of the fertility effect from zero, we conclude that the number of siblings is not a mediator in the relationship of interest.

[29]Using a sample of mothers and the much larger German Microcensus, Piopiunik (2014) finds an effect similar to the one we find, which is significant at the ten percent level of significance, and concludes that the CSL reform improved the probability to have a vocational degree among mothers only marginally.

[30]The calculation of the confidence interval is based on the conventional standard error.

2.6 Conclusion

The incidence and prevalence of mental health disorders is increasing globally (Bloom et al., 2011). With contributions suggesting a substantial intergenerational transmission of mental health status (Johnston et al., 2013), this trend could result in negative long-run consequences, even affecting the next generation. In this chapter, we analyze whether maternal education has an effect on the mental health status of their children in adulthood.

Using exogenous variation in education induced by a CSL reform in Germany, we find that the additional year of maternal schooling does not have any effect on her children's mental health in adulthood. However, we find, consistent with the literature, that maternal years of schooling indeed has a protective effect on the children's physical health. We show that this effect operates mainly through physical functioning. Possible explanations for the absence of the effect on mental health could be the fact, that the CSL reform did not change the school certificate the mothers received. Since the German labor market is a highly formalized labor market, this reform possibly did not lead to additional resources, e.g. higher income or an alternative social network. Alternative empirical strategies that test the relationship of interest at different educational margins could be an interesting venue for future research.

In the second part, we explore potential mediators. We find that the number of close friends is such a mediator. This is a result that is new in the literature. However, the number of close friends mediates the relationship only partially. One reason for this could be the fact that it is also the quality of the friendship that matters.

In summary, while educational interventions show a multitude of positive effects, we find no evidence for positive spillovers of maternal education on offspring's mental health. This adds to the existing - at best ambiguous - findings on the effect of education on one's own mental health suggesting that the predictions of the Grossman model about education and efficiency in health production neither extent to one's own nor to the next generation's mental health.

Appendix

2.A Additional tables

Table 2.A.1: Effect on child's outcomes

	(1) 2SLS estimate	(2) F-stat.	(3) Observations
Panel A: BMI and obesity			
BMI	0.205 (0.459) [0.657]	25.576	3314
Obese	0.010 (0.033) [0.737]		
Panel B: Educational attainment			
Years of schooling	0.114 (0.162) [0.675]	30.681	2897
Intermediate or high school degree	0.046 (0.035) [0.362]		
Panel C: Number of close friends			
Number of close friends	0.733** (0.303) [0.050]	64.168	2287

Note: Table 2.A.1 displays 2SLS estimates for a regression of the child's outcome on maternal years of schooling, instrumented by maternal reform exposure. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth, a state of schooling specific linear trend in the maternal year of birth, a second order polynomial in the child's age, indicators for the each survey year and an indicator for the child being female. First stages statistics are shown only once for specifications that rely on transformations of the same outcome. Equal tail p-values from the Wild-Cluster Bootstrap procedure, with 999 repetitions and Rademacher weights, are in squared brackets. Significance stars are based on p-values based on non-bootstrapped and clustered standard errors and read: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.2: Effects on home environment

	(1) 2SLS estimate	(2) F-Stat.	(3) Observations
Panel A: Father's educational attainment			
Years of schooling	0.638***		
	(0.161)		
	[0.026]	33.401	1753
Higher than basic school-leaving degree	0.131**		
	(0.060)		
	[0.100]		
Panel B: Vocational degree of mother			
Vocational degree	0.024	27.628	2124
	(0.035)		
	[0.579]		
Panel C: Number of children			
Number of children	0.150*	25.215	2021
	(0.086)		
	[0.104]		

Note: Table 2.A.2 displays IV estimates for a regression of maternal outcomes on maternal years of schooling, instrumented by maternal reform exposure. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth and a state of schooling specific linear trend in the maternal year of birth. First stages statistics are shown only once for specifications that rely on transformations of the same outcome. Robust standard errors, in parentheses, are clustered on the maternal state of schooling level. Equal tail p-values from the Wild cluster bootstrap-t procedure, with 999 repetitions and Rademacher weights, are displayed in squared brackets. Robust standard errors, clustered on the maternal state of schooling level, are in parantheses. Significance stars are based on conventional p-values and read: * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.A.3: Association of mediator with MCS score

	(1) Coefficient	(2) F-stat.	(3) Observations
Panel A: BMI and obesity			
BMI	-0.003		
	(0.007)	29.423	3314
Obese	-0.102		
	(0.073)		
Panel B: Educational attainment			
Years of schooling	0.017		
	(0.042)	59.786	2897
Intermediate or high school degree	0.047		
	(0.115)		
Panel C: Number of close friends			
Number of close friends	0.012***	59.311	2287
	(0.004)		
Panel D: Father's educational attainment			
Years of schooling	0.029		
	(0.053)	33.531	2818
Intermediate or high school degree	0.061		
	(0.177)		
Panel E: Vocational degree of mother			
Vocational degree of mother	0.031	33.537	3316
	(0.116)		
Panel F: Number of children			
Number of children	0.040***	42.310	3158
	(0.010)		

Note: Table 2.A.3 displays coefficients on the respective mediating variables in an IV regression of the child's MCS score on maternal years of schooling, instrumented by maternal reform exposure. Each regression includes indicators for maternal state of schooling, indicators for maternal year of birth, a state of schooling specific linear trend in the maternal year of birth, an indicator which is equal to one if the child is female, a second order polynomial in the child's age and survey year indicators. First stages statistics are shown only once for specifications that rely on transformations of the same outcome. Robust standard errors, in parantheses, are clustered on the county level. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

2.B Additional figures

105. How would you describe your current health?

- Very good ☐
- Good ☐
- Satisfactory ☐
- Poor ☐
- Bad ☐

106. When you have to climb several flights of stairs on foot, does your health limit you greatly, somewhat, or not at all?

- Greatly ☐
- Somewhat ☐
- Not at all ☐

107. And what about other demanding everyday activities, such as when you have to lift something heavy or do something requiring physical mobility: Does your health limit you greatly, somewhat, or not at all?

- Greatly ☐
- Somewhat ☐
- Not at all ☐

108. During the last four weeks, how often did you:	**Always**	**Often**	**Some-times**	**Almost never**	**Never**
• feel rushed or pressed for time?	☐	☐	☐	☐	☐
• feel down and gloomy?	☐	☐	☐	☐	☐
• feel calm and relaxed?	☐	☐	☐	☐	☐
• feel energetic?	☐	☐	☐	☐	☐
• have severe physical pain?	☐	☐	☐	☐	☐
• feel that due to physical health problems					
– you achieved less than you wanted to at work or in everyday activities?	☐	☐	☐	☐	☐
– you were limited in some way at work or in everyday activities?	☐	☐	☐	☐	☐
• feel that due to mental health or emotional problems					
– you achieved less than you wanted to at work or in everyday activities?	☐	☐	☐	☐	☐
– you carried out your work or everyday tasks less thoroughly than usual?	☐	☐	☐	☐	☐
• feel that due to physical or mental health problems you were limited socially, that is, in contact with friends, acquaintances, or relatives?	☐	☐	☐	☐	☐

Figure 2.B.1: SF-12 questionnaire of the SOEP.

Note: Source: TNS Infratest Sozialforschung (2014).

CHAPTER 3

Hate is too great a burden to bear: Hate crimes and the mental health of refugees*

Against a background of increasing violence against non-natives, we estimate the effect of hate crime on refugees' mental health in Germany. For this purpose, we combine two datasets: administrative records on xenophobic crime against refugee shelters by the Federal Criminal Office and the IAB-BAMF-SOEP Survey of Refugees. We apply a regression discontinuity design in time to estimate the effect of interest. Our results indicate that hate crime has a substantial negative effect on several mental health indicators, including the Mental Component Summary score and the Patient Health Questionnaire-4 score. The effects are stronger for refugees with closer geographic proximity to the focal hate crime and refugees with low country-specific human capital. While the estimated effect is only transitory, we argue that negative mental health shocks during the critical period after arrival have important long-term consequences.

*This chapter is joint work with Felicitas Schikora. This chapter has been published as "Hate is too great a burden to bear: Hate crimes and the mental health of refugees." SOEPpaper on Multidisciplinary Panel Data Research 1130 and submitted to the *The Review of Economics and Statistics*.

3.1 Introduction

In the 2010s, the world witnessed two global phenomena: First, forced migration increased dramatically. The number of displaced persons almost doubled from about 42 million in 2008 to 75 million in 2018 (UNHCR, 2019). Second, the prevalence of hate crimes increased markedly. For example, CSTE (2017) reports that hate crime rose by 22 percent in the United States' six largest cities between 2016 and 2017.[1] This marks the third consecutive annual increase for the U.S., a pattern that has not been observed since 2004. We further observe an immense surge of violence against immigrants in Europe (Council of Europe, 2016). The arrival of about 800,000 refugees in Germany in 2015 was accompanied by a sudden increase in hate crimes against refugees (+ 400 percent, BKA).[2]

Hate crimes affect economic behavior through increased feelings of uncertainty, fear, and risk (Becker and Rubinstein, 2011). As a consequence, being victimized is associated with considerable costs in the economic, behavioral, and health domain (Bindler et al., 2020). For example, the costs of victimization amount to two to six percent of the gross domestic product in the U.S. (Chalfin, 2015). One important group which is regularly targeted by hate crimes are migrants, including refugees.

In the economic literature on migration, refugees are considered "permanent" migrants (Dustmann and Glitz, 2011). They remain in their destination country for a long period of time, unable or unwilling to return to their home country, where they are at risk of persecution or conflict. Given that permanent migrants can expect to accrue the returns to integration over longer time horizons than temporary migrants, their lifetime utility strongly therefore depends on their initial integration success (Dustmann and Glitz, 2011).

For this reason, the potential consequences of adverse experiences due to hate crime are particularly consequential for refugees. Therefore, we answer the important question, what are the mental health costs of hate crimes for refugees?

Clearly, mental health shocks can have very detrimental long-term consequences for the victims of hate crimes. For instance, research on the psychological foundations

[1] The increase for all thirteen surveyed cities was 19.9 percent for the period under consideration.

[2] A hate crime is defined as a crime against a specific group of individuals. Typically, hate crimes are committed because of the victim's race, gender, sexuality, color or ancestry (Gale et al., 2002).

of poverty stresses that reduced mental bandwidth increases the likelihood of worse economic choices. Worse economic choices in turn reduce mental bandwidth, resulting in a downward spiral (Schilbach et al., 2016). Similarly, we propose that hate crimes cause mental stress, which in turn may reduce refugees' mental bandwidth. This reduction in their mental bandwidth could impair refugees' economic decision-making ability. This could be particularly detrimental for refugees who fled severe conditions and are at the start of a life in a new country. In addition, a broad literature shows the adverse consequences of childhood exposure to stress and adverse conditions, including in-utero exposure to severe stress, on an individual's long-term life outcomes (Almond et al., 2018; Almond and Currie, 2011). This possibly impairs the life-trajectories of the next generation.[3]

To the best of our knowledge, the effect of hate crime on refugees' mental health has not been assessed in the existing literature. There are two reasons for this short-coming: First, we need data that combines both, representative information on refugees' mental health and their place of residence as well as information on a wide range of individual characteristics. Second, unobservable variables potentially bias the relationship between the occurrence of a hate crime and refugees' mental health. For instance, refugees may choose their place of residence endogenously based on regional characteristics, such as favorable economic conditions or existing ethnic networks, which may jointly determine both refugees' mental health and the occurrence of hate crime. Thus, it is essential to rely on an identification strategy that allows for the consistent estimation of the effect of hate crime on refugees' mental health. We advance the literature by solving these two problems.

To estimate the effect of hate crime on refugees' mental health, we rely on a regression discontinuity in time (RD) design (Hausman and Rapson, 2018).[4] Using German counties that experience at least one hate crime against a refugee shelter, we assign each refugee the closest hate crime in the respective county measured in days elapsed since this focal hate crime. We then compare refugees' mental health immediately before and after an attack on the county level. Thus, the identification of our effect

[3]For instance, Persson and Rossin-Slater (2018) show that prenatal exposure to stress increases take-up of ADHD medications during childhood and take up of depression medication later in life. Further, the infants' indirect in-utero exposure to the 9/11 attacks in the U.S. caused their birth weight to decrease by 15 grams, the likelihood of being born weighting less than 1,500 grams by 14%, and the likelihood of being born at less than 37 gestational weeks by 9% (Brown, 2020).

[4]In what follows, we refer to our research design as RD design.

relies on the assumption that refugees' mental health is a continuous function of the number of elapsed days since the focal hate crime. We find strong support for this assumption, emphasizing the credibility of our research design.

Our empirical analysis relies on the unique IAB-BAMF-SOEP Survey of Refugees in Germany as well as geo-referenced administrative data on hate crimes from the Federal Criminal Office (BKA). The IAB-BAMF-SOEP Survey of Refugees is a representative survey of refugees who arrived in Germany between 2013 and 2016.[5] The data provides information on refugees' migration histories, background characteristics as well as overall living conditions and integration outcomes. Most importantly, it includes information on the exact interview date, the place of residence, and high-quality information on refugees' mental health. Our two mental health measures included in the IAB-BAMF-SOEP Survey of Refugees are the Mental Component Summary (MCS) score and the Patient Health Questionnaire-4 (PHQ-4) score. These are two well-established summary measures of general mental health as well as anxiety and depression, respectively. In order to link our analysis to related studies, including Deole (2019) and Steinhardt (2018), we also investigate the effect of hate crime on refugees' life satisfaction or intention to stay at the extensive margin.

Our second source of information is the BKA data, which reports hate crimes against refugee shelters. The BKA data contains time, place, the type of crime, and the crime's political motivation. This allows us to geo-reference the information and combine the administrative data on hate crimes with the IAB-BAMF-SOEP Survey of Refugees. The advantage of the administrative BKA data is that it contains information on hate crime directed toward refugees' shelters, which unambiguously represent hate crimes. This is an advantage over other data sources that do not differentiate between hate crime directed toward refugees or other residents with a migration background. Thus, we focus on refugee shelters since these are very salient forms of hate crime. In addition, data from non-administrative sources, such as newspapers, could suffer from endogenous coverage (Entorf and Lange, 2019).

Our results indicate that the experience of a hate crime reduces refugees' MCS score by 37% of a standard deviation. Similarly, hate crimes reduce refugees' PHQ-4

[5]The IAB-BAMF-SOEP Sample of Refugees in Germany is part of the German Socio-Economic Panel (SOEP). We use version 34 of the SOEP. DOI: 10.5684/soep.v34.

score by 28% of a standard deviation.[6] In contrast to existing studies that focus on economic migrants in Germany, such as Deole (2019) and Steinhardt (2018), we find no effect on refugees' life satisfaction. A potential reason for this may be the fact that refugees draw from different segments of the population in their home country than do economic migrants.[7] For instance, Deole (2019) and Steinhardt (2018) focus on the population of migrants who moved to Germany in the late 1960s to meet the shortage of labor that was prevalent in Germany at that time. These migrants were actively recruited, either by the German government or the sending countries' government. Furthermore, we find no effect on refugees' intention to stay (ITS) in Germany at the extensive margin. This is an important result. Existing research shows that the time horizon over which migrants can accrue returns to investments in country-specific human capital is positively associated with the gradient in their age-earnings profile (e.g., Dustmann, 2000, 1993, 1997). With hate crimes having little effect on the refugees' ITS, we conclude that a change in the ITS can be ruled out as a mediator between hate crime and the accumulation of country-specific human capital.

We also find strong suggestive evidence that our effects are mostly driven by refugees living in close proximity to the focal hate crime. Our data allows us to calculate geographical distances between the location of the focal hate crime, e.g., city, town or municipality, and the refugees' place of residence. We then perform a median split distinguishing between refugees who live close to the hate crime and refugees who live further away. We find that refugees living closer to the respective hate crime have also stronger adverse mental health effects, while the effects are considerably smaller and insignificant for refugees living further away. Thus, this finding shows that the mental health effects reflect a response to a more direct exposure to hate crime.

In a second part, we test Becker and Rubinstein's (2011) conjecture that individuals with higher cognitive abilities are more likely to overcome the shock caused by hate crimes. To be more precise, Becker and Rubinstein's (2011) mechanism suggests that individuals with higher ability are more likely to align the objective and subjective likelihood of becoming a victim of a hate crime. For migrants, such as refugees, we

[6]Typically, the PHQ-4 score indicates the intensity of symptoms of depression and anxiety. To allow for a consistent comparison with the MCS score, we inverted the scale. Thus, a higher score indicates better mental health.

[7]Economic migrants normally leave their country of origin because of pull rather than push factors.

contend that it is the country-specific human capital, i.e. language proficiency, that helps refugees to assess the true risk of being harmed by hate crimes. Therefore, we show that refugees, who are better integrated within their host country's society–e.g., those who have frequent contact with German natives or possess higher language proficiency levels–are less severely affected if they experienced a hate crime. This effect is most prevalent for the PHQ-4 score: While the estimated effect amounts to roughly 45% of a standard deviation for respondents who report low levels of German language proficiency, the effect size is halved for refugees with high country-specific human capital and statistically insignificant. Hence, in line with Becker and Rubinstein (2011), we find that individuals, who are more likely to align the subjective with the objective probability of being harmed by a hate crime, are less severely affected by hate crimes. In addition, our analysis suggests an important role for the refugees' ability to acquire information about these hate crimes. Lastly, our test, in which we interact country-specific human capital with the opportunity costs of acquiring this human capital, shows that "general ability" and the stock of country-specific human capital are complementary.

Moreover, while our empirical results indicate that the effect dissipates after approximately three months, we argue, similar to Clark et al. (2020), that those shocks have considerable long-term consequences via the reduced mental bandwidth, which can lead to worse economic decisions and thus, detrimental long-term consequences (Becker and Rubinstein, 2011; Schilbach et al., 2016), potentially also affecting the next generation (Almond et al., 2018; Almond and Currie, 2011).

This chapter relates to five branches of the literature. Abstaining from immigration, previous papers unanimously conclude that terrorist attacks have substantial negative effects on individuals' life satisfaction that persist, albeit, only temporarily (Clark et al., 2020; Akay et al., 2020). Using the 9/11 terrorist attacks as a quasi-experiment, Metcalfe et al. (2011) further shows that there are spillover effects to other countries such as the U.K..

Second, we also contribute to the literature on the effect of hate crimes on immigrants' health and integration within the host society. For the U.S., Gould and Klor (2014) show that the 9/11 attacks induced a backlash against Muslim immigrants, which in turn increased the opportunity costs of assimilation. For instance, in response to the 9/11 attack, Muslim immigrants in the U.S. were more likely to marry someone with the same ethnic background than before. Further, they also experienced lower rates of labor force participation (Gould and Klor, 2014). For Germany, there is evidence that hate crimes reduce integration outcomes as well as life and health satisfaction for immigrants with a Turkish background. Steinhardt (2018) shows that macro exposure to anti-immigrant attacks in the early 1990s in West Germany reduces the Turkish migrants' life satisfaction, increases their return intentions, and slows language acquisition. Further, Deole (2019) studies the revelation crimes directed toward Turkish residents in Germany in 2011. Deole (2019) finds that these revelations reduced the Turkish immigrants' life satisfaction.

We also relate to the literature focusing on the causes of hate crime. For Germany, Krueger and Pischke (1997), Falk et al. (2011) and Entorf and Lange (2019) analyze which socio-demographic characteristics predict hate crimes on the county level. Moreover, the literature is investigating how social media can predict hate crime (Bursztyn et al., 2019; Müller and Schwarz, 2020; Müller and Schwarz, 2020). We add to this literature by turning to the effect of hate crime on the most vulnerable group among those targeted: refugees. Lastly, we also contribute to a larger more general literature about the socioeconomic determinants of mental health (e.g. Adhvaryu et al., 2019; Fruewirth et al., 2019).

Our contribution to the literature is twofold: First, to the best of our knowledge, we are the first to analyze the effect of hate crime on refugees' mental health. This is surprising, given the stark increase in forced migration, which is expected to increase further given the economic and environmental changes worldwide (UNHCR, 2019), and the fact that mental illnesses has the highest prevalence of all non-communicable diseases (Bloom et al., 2011). Our results further suggest the importance of mental health for (labor market) integration and the subsequent long-term consequences for refugees in Germany.

Second, we test the importance of the refugees' opportunity to acquire information as a mediator between hate crime and the refugees' mental health response. This allows to further characterize Becker and Rubinstein's (2011) prediction that individuals with higher ability are more likely to overcome the shock due to hate crimes.

3.2 Forced migration and hate crime

In the 2010s, environmental deterioration and political upheavals in many African and Asian countries caused a stark increase in the number of refugees worldwide.

Figure 3.B.1 in the appendix shows that the trend accelerated starting in 2013, following the outbreak of the Arab spring. Among refugees, the vast majority typically migrates either within their country of origin or settles down in a neighboring country (UNHCR, 2019). However, as the supply conditions deteriorated rapidly in the neighboring countries' refugee camps and intermediary states like Libya collapsed, large numbers of refugees began to migrate to Central Europe in 2014 and 2015 (Luft, 2016).

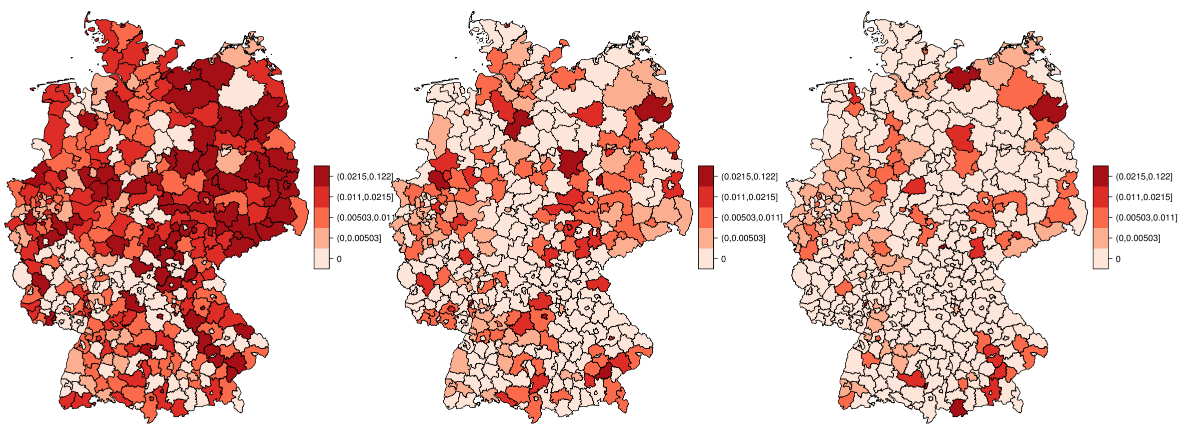
In Europe, the Dublin regulation stipulates that an application for asylum must be processed by the first Dublin country the asylum seeker enters. Therefore, European Union (EU) members closest to the refugees' countries of origin–normally at the edge of the EU–were disproportionately affected by the number of refugees migrating to Europe. As the number of refugees increased in these countries, the local conditions deteriorated quickly. Initially, the European countries tried to negotiate a new scheme to distribute refugees across the European Union's member countries. However, these negotiations were unsuccessful and, finally, in light of the inhumane situation of the refugees in some of the EU's host countries, the German government suspended the Dublin regulation in fall 2015 (BAMF, 2015). This triggered a large influx of refugees to Germany. Consequently, in 2015 Germany received the largest number of refugees in absolute terms, ranking third after Austria and Sweden in relative terms (Organization for Economic Co-operation and Development, 2017). Subsequently, however, the number of refugees in Germany decreased to pre-2015 levels (Figure 3.B.2 in the appendix).

Turning to the refugees' demographics, the majority of refugees in Germany originate from Syria, Afghanistan, and Iraq. In 2016, the Federal Office for Migration and Refugees (BAMF) reported the share of first-time asylum applications was 36.9% Syrian, 17.6% Afghan, and 13.3% Iraqi (BAMF, 2016). In addition, these refugees tend to be very young with 73.8% of these refugees younger than 30 years of age (BAMF, 2016).

Associated with the stark increase in the number of refugees, Germany experienced a strong increase in xenophobic sentiments directed against immigrants and refugees. For instance, using data from the Federal Criminal Office, we observe a strong increase in hate crimes against refugee shelters around the time when large number of refugees entered Germany (Figure 3.B.3 in the appendix). The number of these hate crimes increased strongly from 2014 to 2015, remained on an elevated level in 2016, and then returned to initial levels as the number of foreigners arriving in Germany fell. For instance, while our data shows 971 hate crimes in 2016, it declines to 303 hate crimes in 2017 and 170 in 2018.

Figure 3.1 provides a more detailed picture, displaying the number of attacks on refugee shelters per 100,000 residents at county-level per year.[8] We make two observations from this figure. First, as described before, the intensity of hate crimes declines over time. Second, although hate crime is always more prevalent in Eastern German states, it is also widely dispersed across Germany.

Figure 3.1: Number of attacks on refugee shelters per 100,000 inhabitants and counties



(a) 2016 (b) 2017 (c) 2018

Note: Figures 3.1a to 3.1c display the number of attacks on refugee shelters per 100,000 inhabitants and county from 2016 until 2018, respectively. Source: BKA data.

[8]We resort to hate crimes per capita in these figures since the initial distribution of refugees within states relies on the counties' share of the population within each state. Consequently, a cross-sectional regression of the counties' share of the states' intake of refugees on the counties' population share within the respective state and state fixed effects results in an estimated OLS coefficient of one. Results are available on request.

3.3 Data

We use two innovative datasets to estimate the effect of hate crime on refugees' mental health: The first dataset incorporates administrative information on hate crime against refugee shelters. The second dataset is the IAB-BAMF-SOEP Survey of Refugees in Germany, which provides us with detailed information on refugees' mental health as well as a wide range of socio-economic characteristics. In what follows, we describe these two datasets in detail.

3.3.1 Administrative data on hate crime against refugee shelters

Our comprehensive data on hate crime against refugee shelters stems from the German Federal Criminal Police Office ("*Bundeskriminalamt*"). The information was compiled by the German Federal Government in response to small inquiries ("*kleine Anfrage*") of the parliamentary group "DIE LINKE" and is published on a quarterly basis (e.g. Bundestag, 2016). Each entry in these files comprises information on the date of the attack, the state, the locality, the type of crime, and the crime's political motivation. For illustrative purposes, Table 3.A.1 in the appendix illustrates an excerpt of the data for January 1, 2016. The major advantage of this dataset is that it reports hate crimes that target refugees specifically rather than an aggregate measure on hate crimes that would have precluded the ability to distinguish between economic migrants and refugees. Second, hate crime against refugee shelters are much more salient than individual incidents such as refugees being attacked on the street. Finally, the BKA data is less likely to suffer from endogenous coverage, which could, for instance, be the case for newspaper data (Entorf and Lange, 2019). As such, it is an ideal source of information on hate crime against refugees for this analysis.

In a first step, we collected all information on hate crime against refugee shelters from the small inquiries and digitized the BKA information accordingly. In a second step, we geo-referenced the data based on information on the state and the exact location, e.g., the name of the city or municipality.[9] Overall, our data records 1,444

[9]In less than five cases, we were not able to determine the exact GPS location since the respective location existed several times in the respective state.

events between 2016 and 2018. As displayed in Section 3.2, the incidence of hate crimes against refugee shelters has substantially decreased over time.

3.3.2 IAB-BAMF-SOEP Survey of Refugees

The IAB-BAMF-SOEP Survey of Refugees comprises information on refugees' mental health and their socio-economic characteristics. The IAB-BAMF-SOEP Survey of Refugees has been introduced in 2016, in response to the major influx of migrants to Germany in 2015 (Brücker et al., 2016; Deutsches Institut für Wirtschaftsforschung, 2017). This novel survey is part of the Socio-Economic Panel (SOEP) (Göbel et al., 2018) and, hitherto, is the only data base that allows for quantitative and empirical social research on this timely manner. Besides information on refugees' migration histories, background characteristics, overall living conditions, and integration outcomes in Germany, the IAB-BAMF-SOEP Survey of Refugees provides detailed information on refugees' mental health and their exact place of residence.

In our analysis, we make use of a single cross-section from the year 2016 due to two reasons: First, 2016 is the year of the last decade in which hate crimes were most prevalent in Germany. Second, information on the PHQ-4 score is only available in 2016. The interviews in the IAB-BAMF-SOEP Survey of Refugees typically take place from June to December in each year. Consequently, our period of observation is the second half of 2016. We merge each observation in the IAB-BAMF-SOEP Survey of Refugees to the respective hate crimes, based on the information on the exact interview date and the location. For each survey respondent, we then calculate the number of elapsed days since the most recent hate crime–the focal hate crime–in the county of residence. This running variable then governs the treatment status. The running variable is negative for refugees who have been interviewed before the focal hate crime occurred. If, on the other hand, the focal hate crime took place before the refugee was interviewed, the running variable is positive thereby marking the respondent as a treated individual.

3.3.3 Measuring refugees' mental health

We measure the refugees' mental health by the two mental health measures available in the IAB-BAMF-SOEP Survey of Refugees: the MCS and the PHQ-4 score. The

MCS score is based on the Short Form-12 (SF-12) questionnaire, which contains twelve health-related items inferring the respondent's physical and mental health within 30 days preceding the interview (Andersen et al., 2007). The MCS score has been shown to be highly predictive for mental illnesses in the European population (Vilagut et al., 2013) and is an established measure of mental health in the economic literature (e.g. Marcus, 2013; Eibich, 2015; Hofmann and Mühlenweg, 2018).[10]

For the principal component analysis, we combine the twelve health items in eight subscales and normalize these subscales to have mean zero and standard deviation one. Subsequently, we perform a principal component analysis of these eight subscales for all first-time respondents in 2016 and 2017. The eight subscales of the SF-12 questionnaire load exactly on two factors. Figure 3.B.4 in the appendix, which plots the factors against the respective Eigenvalues, shows that the first two factors have Eigenvalues greater or equal to one. We conclude that the first two factors are the only significant factors. In a last step, we perform a varimax rotation. The resulting factor loadings are displayed in Table 3.A.2 in the appendix.

Clearly, the factor loadings of the second factor in column (2) of Table 3.A.2 in the appendix load very high on the subscales that are associated with mental health. The respective factor loadings for the mental health subscales range from 0.577 to 0.823, whereas the remainder factor loadings range from 0.084 to 0.313. In what follows, we refer to this factor as the MCS score.

Along with the MCS score, we also employ a mental health measure based on the PHQ-4 inventory (Kroenke et al., 2009). The scores based on the PHQ-4 inventory have been shown to have high reliability and validity (e.g. Kroenke et al., 2009; Loewe et al., 2010) and, importantly, to have good psychometric properties in a representative survey of Arab refugees (Kliem et al., 2016). The PHQ-4 inventory consists of four items, including the frequency of feeling little interest or pleasure in one's activities, melancholy, anxiety, and the inability to stop worrying. Responses to the four items are given on a four-point Likert-scale, ranging from one "Not at all" to four "(Almost) every day". In what follows, we proceed similar to the construction of the MCS score and perform a principal component analysis of the PHQ-4 inventory.[11] Figure 3.B.5

[10]We apply the algorithm of Andersen et al. (2007) to the IAB-BAMF-SOEP Survey of Refugees. The number of factors as well as the factor loadings are very similar to those of the SOEP norm population in Andersen et al. (2007).

[11]Often, researchers just use the sum of the four items, implying equal weighting of each factor.

in the appendix shows that the Eigenvalue of the first factor is 2.40. In contrast, the second eigenvalue is 0.76. Consequently, we use the first factor as the only significant factor. Additionally, the factor loadings of the first and only factor, depicted in Table 3.A.3 in the appendix, range from 0.598 to 0.845. We label this factor PHQ-4 score. Initially, higher scores indicate worse mental health. However, to ease interpretation, we invert the scale. In this study, higher values are indicative of better mental health.

3.3.4 Additional outcomes and covariates

Additional outcome variables are life satisfaction and the respondents' intention to stay in Germany. Life satisfaction is inferred by the answer to the question "How satisfied are you with your life, all things considered?". The answers to this question are given on an eleven-point Likert-scale, ranging from zero, "Completely dissatisfied", to ten, "Completely satisfied". The respondents' intention to stay is inferred from the answer to the question "Do you want to stay in Germany forever?" Based on responses to this item, we construct an indicator which is equal to one if a respondent wants to stay in Germany forever and zero otherwise.

Additionally, we use the command over the German language as well as the number of contacts with Germans as proxies for country-specific human capital. The respondents are asked how well they can speak, read, or write in German. Answers are given on a five-point Likert-scale ranging from one "Very well" to five "Not at all". We construct an indicator which is equal to one if respondents state that they can speak, read, or write German at least averagely. The time spent with Germans is inferred by a six-point Likert-scale that ranges from one "Daily" to six "Never", with three "Weekly" being the median category. We construct an indicator which is equal to one if a refugee states that he or she has at least weekly contact with Germans. The final summary characteristics, together with further predetermined characteristics, of our working sample are displayed in Table 3.1.[12]

However, we decided to use an equal procedure as with the SF-12 questionnaire to remain consistent across mental health measurements.

[12]In Table 3.1, SIA is the acronym for "Syria, Iraq of Afghanistan". These are the countries from which most of the refugees in the data come from.

Table 3.1: Summary statistics

	Mean (1)	S.D. (2)	Min. (3)	Max. (4)	N (5)
Outcomes:					
MCS score	49.357	10.471	15.075	74.147	1215
PHQ-4 score	0.010	1.031	-1.098	3.150	1215
Life satisfaction	7.280	2.326	0.000	10.000	1215
Intention to stay	0.947	0.225	0.000	1.000	1215
Refugee's characteristics:					
Female	0.388	0.488	0.000	1.000	1215
Year of birth	1981.821	10.376	1940.000	1998.000	1215
Country origin SIA	0.769	0.422	0.000	1.000	1215
Child present	0.674	0.469	0.000	1.000	1215
Married	0.686	0.464	0.000	1.000	1215
Characteristics of counties (2014):					
GDP per capita (in 1000 Euro)	35.939	12.876	20.373	93.773	1215
Average age	44.644	1.801	41.100	49.700	1215
Share of foreigners	0.085	0.050	0.013	0.240	1215

Note: Table 3.1 displays summary statistics for our outcomes, refugees characteristics and the characteristics of the counties in 2014. Column (1) displays means. Column (2) displays the corresponding standard deviations. Column (3) and (4) display the minimum and the maximum. The sample is restricted to a bandwidth of 90 days around the cutoff. Source: SOEP, v34.

3.4 Empirical method

We estimate the effect of hate crime on refugees' mental health using a RD design that compares refugees who have been interviewed shortly before and after a hate crime in the respective county of residence occurred. Thus, we estimate the following weighted local linear regression:

$$Y_{icmd} = \alpha + \beta D_{icmd} + \gamma Dist_{icmd} + \delta D_{icmd} \times Days_{icmd} + \zeta month_i + \theta dow_i + \epsilon_{icmd}. \tag{3.1}$$

In Equation 3.1, Y_{icmd} is the mental health outcome of interest, i.e., the MCS or the PHQ-4 score for respondent i in county c, in month m and day of week d. The indicator D_{icmd} is equal to one if the refugee was interviewed after a hate crime happened in the county of residence and zero otherwise. The running variable $Days_{icmd}$ captures the number of days elapsed since the focal hate crime occurred. We allow for differential linear trends before and after the focal hate crime. Consequently, we include the interaction term $D_{icmd} \times Days_{icmd}$ in Equation 3.1.[13] In addition, we account for potential seasonality in the mental health outcomes by including indicators for the month when respondents were interviewed, $month_i$. Further, we account for potential discontinuities in mental health and the likelihood that a hate crime takes place, which are associated with the day of week, dow_i. For instance, perpetrators could be more active on weekends than on weekdays. At the same time, refugees' mental health could be better on weekends compared to weekdays. In this case, we potentially underestimate the true effect of hate crimes. The inclusion of day of week indicators helps to account for this. We use a triangular kernel and cluster the standard errors on the running variable level, because our running variable is discrete (Lee and Card, 2010).[14]

It is notable that in some counties, hate crimes are clustered in time. Thus, it could be the case that refugees in the control group are treated if they were subject to a hate crime which took place before the focal hate crime, e.g., if a hate crime happened before an individual was observed and the number of days between this other hate crime and the day of observation is at least the number of days until the focal hate crime plus one. Similarly, treated refugees could have been subject to an additional hate crime before the focal hate crime. If this happens randomly, e.g., if these confounding attacks are independent and identically distributed, this would result in an attenuation bias. This attenuation bias potentially causes our estimates to be attenuated towards zero. Therefore, in the robustness section, we carry out a careful test gauging the relevance of this bias. We carefully drop observations that are multiply treated within various bandwidths and observe that the estimates tend to increase as we drop observations that are treated multiple times. Indeed, we find evidence for our conjecture. Thus, as precautionary measure and to optimally utilize the number of observations, we drop refugees who experienced a hate crime within 30 days before the focal hate crime. Based on this empirical specification, we choose the

[13] In the robustness section, we also allow for quadratic trends in the running variable. Our conclusions remain unaltered.

[14] We also base our inference on standard errors clustered on the county level in the robustness section. Our conclusions remain unchanged.

bandwidth to be ±90 days.[15][16]

Continuity assumption Our identification assumption is based on the premise that, in absence of the treatment, the population mean in mental health is a continuous function of the running variable (Hahn et al., 2001). Another way to think about this is by means of selection on observables (Lee and Lemieux, 2010). In our case, the number of days elapsed since the focal hate crime governs the treatment assignment. Refugees who were interviewed before the focal hate crime are part of the control and refugees interviewed after the focal hate crime are considered part of the treatment group. Strictly speaking, the common support in the running variable is not guaranteed in this setting. Therefore, we require the continuity assumption to finally ensure the overlap condition.

Under this assumption, the estimate of γ can be interpreted as the causal effect of hate crime on refugees' mental health. However, we can not directly test the continuity assumption because it involves a counterfactual situation, i.e., we need to observe the population mean through the cutoff in absence of the treatment. Yet, we provide evidence that the continuity assumption holds. If predetermined individual and county level characteristics evolve continuously around the focal hate crime, we may interpret this as empirical evidence that the continuity assumption is valid. Any significant discontinuity in the mental health outcomes around the focal hate crime can be fully attributed to the focal hate crime. To test this, we apply our empirical specification to various predetermined characteristics on the individual and county level.

Our estimates reveal no discontinuity in the predetermined individual and county level characteristics around the focal hate crime. Figure 3.B.6 in the appendix displays RD plots for various predetermined individual and county characteristics. Overall, we find little visual evidence for discontinuities around the focal hate crime. Table 3.2 summarizes the results formally. Column (1) displays the point estimates, column (2) displays the corresponding standard errors, and column (3) displays the p-values associated with the coefficient estimates. Throughout, most of the estimates are small

[15]Following Calonico et al. (2014), we find that the asymptotically MSE-optimal bandwidths for the PHQ-4 and MCS score, life satisfaction are 88.3, 78.1, 71.8 and 112.7, respectively. For expositional clarity, we choose a bandwidth of 90 days or 3 months, which is close to the average of the three respective bandwidths. However, we show that our results are robust to a wide range of bandwidth choices in Section 3.5.3.

[16]We use the Stata package **rdrobust**. For the documentation, please refer to Calonico et al. (2017).

in relative terms. Further, all estimates are statistically insignificant. Thus, we are confident that the continuity assumption is warranted.

Table 3.2: Continuity of predetermined characteristics around the focal hate crime

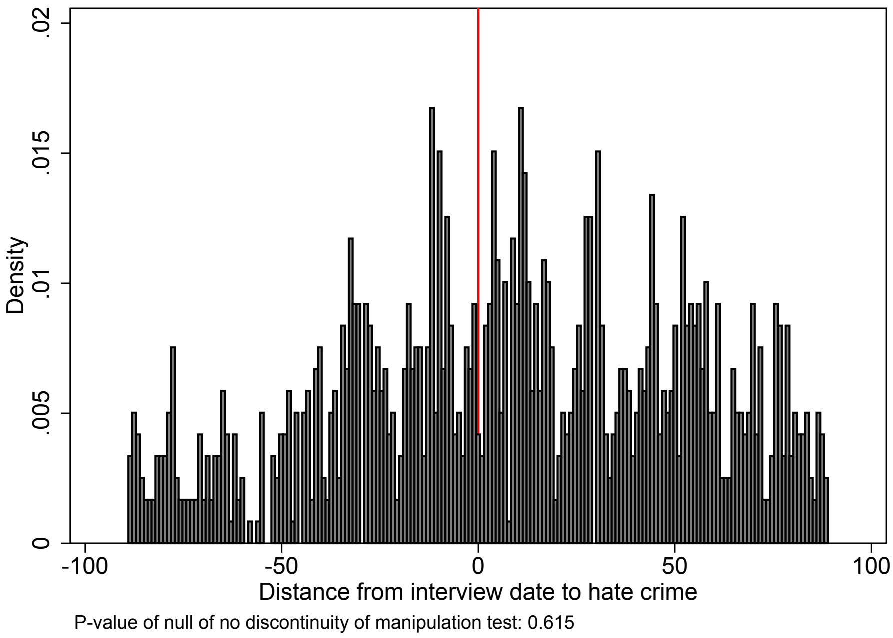
	Point estimate (1)	Standard error (2)	P-value (3)
Child present	-0.081	0.059	0.168
Country of origin SIA	0.050	0.052	0.328
Female	0.007	0.042	0.874
Married	0.033	0.065	0.606
Year of birth	-0.560	1.021	0.583
Average age in county	0.105	0.255	0.679
GDP per capita in county	0.243	1.886	0.897
Share of foreigner in county	-0.004	0.008	0.612

Note: Table 3.2 displays results for a test of the continuity assumption for predetermined individual and county characteristics. Column (1) displays point estimates. Column (2) displays standard errors associated with the point estimates. Column (3) displays p-values. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights and a bandwidth of 90 days. Non-binary outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

Precise manipulation around the cutoff A potential threat to our RD design could be the precise manipulation around the cutoff (Lee and Card, 2010; McCrary, 2008). If selection into or out of the treatment would be possible based on expected gains, our estimate of γ would suffer from selection bias and may be inconsistent. In our context, individuals would desire to select out of treatment. The more vulnerable the refugees are, the more likely they desire to select out of the treatment group. This would bias our estimate of the effect downwards.

Since our data on hate crimes is based on official crime statistics, we assert that strategic manipulation around the cutoff is difficult, if not impossible. This conjecture assumes that these hate crimes are typically not known to the public beforehand. In addition, the SOEP interviews are usually scheduled well in advance. The reason is that the interviews usually take some time, especially if a household consists of multiple individuals. In consequence, it is very unlikely that selection based on expected gains is prevalent.

Figure 3.2: Checking for precise manipulation around the cutoff



Note: Figure 3.2 displays the empirical pdf of observations around the cutoff. A bandwidth of 90 days is choosen. Each bin corresponds to one day. Each par corresponds to the density of observations at each day. The vertical bar indicates the day of the xenophobic attack. The p-value corresponds to a p-value of a manipulation test based on local polynomial regressions of order two. Source: SOEP, v34.

However, if exposure to hate crimes decreases the likelihood that respondents thoroughly reply to all questions of the interview, our estimates would be biased downward since only the most robust respondents would be able to reply. But this results in a testable assumption. If exposure to a hate crime is associated with a lower likelihood that refugees provide information in the SOEP-interviews, we would observe a discontinuity in the empirical distribution of observations around the cutoff.

A density test around the cutoff, that was proposed by McCrary (2008), suggests that neither of the two phenomena are relevant in our case. If individuals were able to select into or out of the treatment or if fewer respondents provided information about their mental health in response to the hate crime, we would detect a discontinuity in the empirical probability density function of interviews around the focal hate crime. Figure 3.2 displays the empirical distribution of observations against the running variable. The vertical line indicates the day of the focal hate crime. Based on the inspection of the empirical probability density function, we find no evidence of a discontinuity around the focal hate crime. A p-value of 0.615 of a formal manipulation test, based on local polynomial regressions of order two (Cattaneo et al., 2020), indicates that there exists no discontinuity around the focal hate crime. Thus, we confidently rule out manipulation or differential response behavior around the cutoff.

3.5 Results

In this section, we report our estimation results as well as additional robustness checks. Thereafter, we report heterogeneity analyses with respect to the refugees' country-specific human capital and the geographic proximity to the focal hate crime.

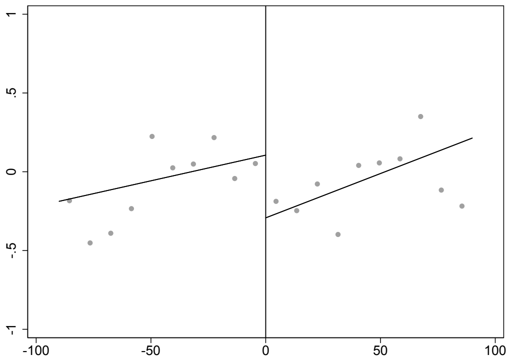
3.5.1 The effect of hate crime on mental health

Figures 3.3a and 3.3b illustrate the main results for the refugees' MCS and PHQ-4 score, whereby each figure displays a local linear fit on either side of the cutoff, with bandwidths of 90 and triangular kernels. We partial out day of week as well as month fixed effects. In both figures, the dots correspond to binned scatterplots, with the number of bins being equal to ten on each side of the cutoff (the focal hate crime). Both mental health outcomes are standardized to have mean zero and standard deviation one.

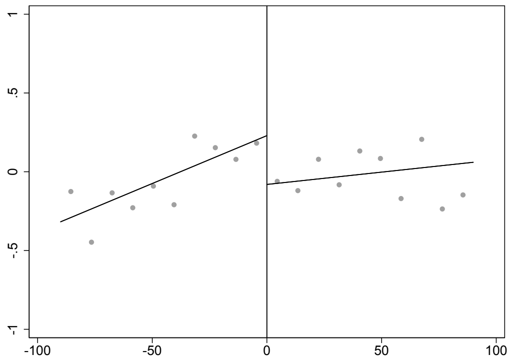
In these descriptive results, we find evidence for a strong discontinuity around the cutoff for the mental health outcomes, suggesting that being a victim to a hate crime worsens the refugees' mental health. Table 3.3 displays effect sizes corresponding to Equation 3.1. Columns (1) and (2) display the point estimates for the MCS and PHQ-4 score along with the standard errors, respectively. Based on these results, the effect sizes correspond to 37% of a standard deviation for the MCS score and 28% of a standard deviation for the PHQ-4 score. As a comparison, Clark et al. (2020) find that the Boston marathon bombing reduced the nearby resident's subjective well-being by a third of a standard deviation. In addition, Metcalfe et al. (2011) find that the 9/11 attack in the U.S. decreases mental distress in the U.K. population by about 7 to 14% of a standard deviation. Thus, our results are of comparable magnitude of studies such as Clark et al. (2020).

While the estimated effects are sizable, the fact that the mental health outcomes are trending toward pre-treatment levels after the hate crime indicates that the ef-

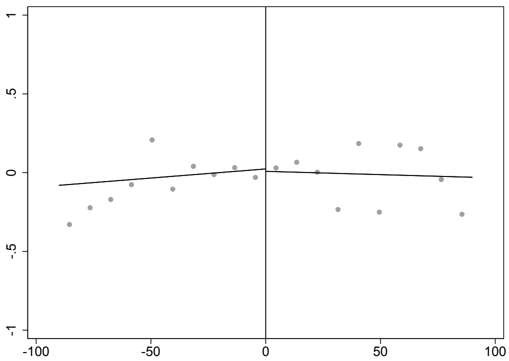
Figure 3.3: Visualization of results



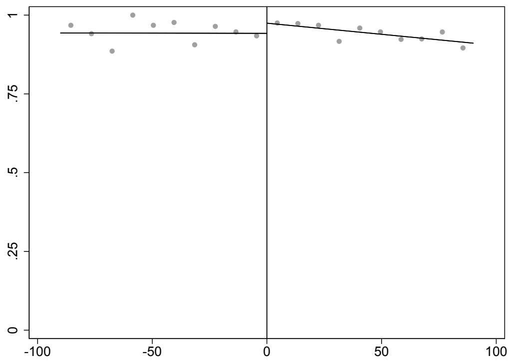
(a) MCS



(b) PHQ-4



(c) Life satisfaction



(d) ITS

Note: Figures 3.3a to 3.3d display the effect of xenophobic attacks on migrants' mental health, life satisfaction and intention to stay. Throughout, the bandwidth is chosen to be 90 days. The dots correspond to a binned scatterplots. The vertical bars are 95% confidence intervals for the means of the bins, based on standard errors that are clustered on the running variable. The linear fit corresponds to a local linear regression with a triangular kernel as in Equation 3.1. Source: SOEP, v34.

fect is transitory. The mental health outcomes reach their pre-treatment level after approximately three months (Figures 3.3a and 3.3b). Yet, the literature shows that such shocks can impair decision making and alter long-term outcomes, especially for the refugees' children. Thus, negative health shocks have the potential to negatively affect the trajectory of refugees, especially since they must navigate through many uncertainties shortly after arrival.

3.5.2 The effect of hate crime on life satisfaction and ITS

Figures 3.3c and 3.3d illustrate the main results for refugees' life satisfaction and ITS. The corresponding estimation results are displayed in columns (3) and (4) of Table 3.3.

Table 3.3: The effect of hate crime on refugees' mental health, life satisfaction and intention to stay

	MCS	PHQ-4	LS	ITS
	(1)	(2)	(3)	(4)
Effect of hate crime	-0.368***	-0.284***	-0.040	0.023
	(0.118)	(0.104)	(0.103)	(0.022)
Number of observations	1215	1215	1215	1215

Note: Table 3.3 displays the effect of hate crimes on refugees' mental health, life satisfaction and intention to stay. Columns (1), (2), (3), and (4) display point estimates and corresponding standard errors for the MCS score, the PHQ-4 score as well as life satisfaction and refugees' intention to stay, respectively. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights and a bandwidth of 90 days. Non-binary outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

The outcome life satisfaction is standardized to have mean zero and standard deviation one. In contrast to mental health, hate crime has no effect on refugees' life satisfaction and their intention to stay in Germany. These are remarkable results, which stand in clear contrast to the findings of Deole (2019) and Steinhardt (2018).

It must be emphasized that life satisfaction is a multi-dimensional concept that measures the overall quality of life. As such, measures of life satisfaction are clearly distinct but associated with measures of symptoms of common mental disorders, e.g., depression (Keyes, 2006). To put it differently, having impaired mental health is not a sufficient condition for decreased life satisfaction. Furthermore, we argue that this emphasizes the difference in the perception of hate crime between refugees and economic migrants. Refugees are typically unable or unwilling to return to their home countries for fear of violent conflict or persecution. This is particularly true for the refugee population in our sample, who were mainly displaced because of civil wars. Thus, strong push factors caused these refugees to search refuge in Europe. In contrast to refugees, the composition of economic migrants is the result of an interaction of pull and push factors (Lazear, 2021). Consequently, a hypothesis consistent with our empirical observations is that the threshold that causes economic migrants to reconsider their time horizon in the host country is lower than for refugees. In addition, Steinhardt (2018) focuses on the intention to return within the next five years. Therefore, our study and that of Steinhardt (2018) compare the return intentions at different margins, e.g. extensive versus intensive margin.

The observation that hate crimes do not alter the refugees' intention to stay has an additional implication: With hate crimes not altering the time horizon over which refugees accrue returns to country-specific human capital, this cannot be considered an important mediator in the relationship between hate crime and long-term economic well-being.

3.5.3 Additional robustness checks

The previous section shows that hate crime has a strong negative effect on refugees' mental health, i.e., the MCS and the PHQ-4 score, and no effect on the refugees' life satisfaction and ITS. In this section, we provide additional robustness checks supporting the credibility of our estimates.

Choice of bandwidth Our results are robust to a wide range of bandwidth choices. Figure 3.B.7a to Figure 3.B.7d in the appendix display the coefficient estimates of Equation 3.1 and associated 95% confidence intervals as a function of the bandwidth for the MCS and PHQ-4 score as well as life satisfaction and ITS. Varying the bandwidth from 10 to 150 days in increments of 10 days, we see that the coefficient estimates for the MCS and the PHQ-4 score are similar to the main results and statistically significant for a wide range of bandwidths surrounding the respective MSE-optimal bandwidth. In contrast, for all bandwidths, point estimates for life satisfaction and the refugees' ITS are close to zero and statistically insignificant.

Inclusion of covariates The results are also robust to the inclusion of a wide set of predetermined covariates on the individual or county level. In Section 3.4, we argue that identification stems from the assumption that–in absence of the focal hate crime–the population mean of our outcome is a continuous function of the running variable. Alternatively, one can also think of the identification stemming from local randomization around the focal hate crime (Lee and Lemieux, 2010; Hausman and Rapson, 2018). We provide evidence for this by including predetermined individual and county level characteristics in Equation 3.1. The results are displayed in Table 3.A.4 in the appendix, where each row displays coefficients and standard errors for another outcome. We subsequently add different covariates to the regression: Column

(1) adds individuals characteristics, column (2) includes only county level characteristics, and column (3) includes both. Throughout, we observe that the coefficients remain remarkable stable and significant. We consider this as evidence that the local randomization was indeed successful.

Clustering of hate crimes in time In Section 3.4, we argue that the clustering of hate crimes within counties in time may bias our results downwards. If individuals are multiply treated, i.e., if they experience several hate crimes preceding the interview, refugees' mental health may decrease, and our estimates would be attenuated. Therefore, we exclude those individuals who experience a second hate crime within a thirty-day window before the interview. One may, however, argue that this is a selective choice. Hence, as a robustness check, we subsequently exclude different time frames and estimate our treatment effect. Table 3.A.5 in the appendix illustrates estimation results. Overall, we find hate crimes to substantially impair refugees' mental health. In line with our argument, the estimates increase in size the more strictly we ban multiple treated. For instance, while the MCS score decreases by 37% of a standard deviation in our baseline specification, this value increases to 71% of a standard deviation if we exclude observations who experience a second hate crime in a ninety day window before the interview.

More flexible specification Furthermore, our results are robust to alternative and more flexible specifications. The trend in the running variable in our main specification is linear. In general, there is no reason to believe that the trend in our running variable is indeed linear. If we misspecify our model, our estimate could be potentially biased. Therefore, we also allow for a quadratic trend in the running variable and follow the recommendation of Gelman and Imbens (2019) to avoid higher order polynomials than order two.[17] Table 3.A.6 in the appendix displays the results for specifications with a quadratic trend in the running variable. For each outcome, we separately calculated asymptotically MSE-optimal bandwidths (Calonico et al., 2014). Again, the effects point toward a sizable negative effect of hate crimes on refugees' mental health. While the effect size for the MCS score remains relatively stable, the effect size for the PHQ-4 score increases from about 28% of a standard

[17] Gelman and Imbens (2019) argument rests on the observation that a RD estimate is the difference of weighted averages of the outcome on the left and the right of the respective cutoff. In case of higher order polynomials larger than two, the odds are high that very high weights are assigned to observations further away from the cutoff.

deviation to 34% of a standard deviation.

Inference In our main specification, we followed the literature and clustered on the level of the running variable. However, we show that our results are robust to clustering on the county level instead. This implies serial correlation of the regressor or error term on the county level. The results are displayed in Table 3.A.7 in the appendix. Table 3.A.7 displays the respective coefficient estimates with the standard errors, clustered on the county level. Again, our conclusion remain unchanged.

Placebo estimates If we assume that the event happened either thirty days before or after the focal hate crime our estimation results become null. Figure 3.B.8 in the appendix shows coefficient estimates for each of these specifications along with the accompanying 95% confidence intervals. Throughout, the coefficients are small and close to zero. Further, the confidence intervals suggest that we cannot reject the absence of any effect.

3.5.4 Interaction with geographical distance

The previous results rely on the county of residence as the relevant geographical unit. We chose this in order to avoid *ad hoc* assumptions about the relevant distance in, for instance, radius matching. However, a natural question that arises is, whether it is actual hearsay or the fact that the refugees are directly affected by the respective hate crime. To further characterize our estimates, we calculate the actual geographic distance between the place where the focal hate crime in the county of residence took place and the refugees' place of residence. For the IAB-BAMF-SOEP Survey of Refugees, the exact geo-location is available within a specialized secure setting at the Research Data Center of the SOEP. Unfortunately, we do not have the exact GPS data of the refugee accommodations that were attacked. We only have the location, e.g., city or municipality. We assign each attack the centroid of the respective location. However, in some cases, it is the district of a city. Thus, some measurement error is associated with the distance calculation. To minimize measurement concerns, we distinguish between refugees living close by and further away by means of a median split.

The results clearly indicate that those refugees living closer to the focal hate crime show greater effects than those living further away. Table 3.4 shows the corresponding estimates. For those refugees closer to the hate crime, the effect sizes are more than twice as large as for those living further away. In addition, the estimates become insignificant for refugees living further away. However, our estimates are not precise enough to formally reject the null hypothesis of no difference between refugees living closer to the focal hate crime and those living further away. Overall, our conclusions are similar to the conclusions drawn by Clark et al. (2020), who also find that residents who live more closely to the Boston Marathon Bombing are more severely affected.

Table 3.4: The effect of xenophobic attacks on refugees' mental health, accounting for geographic distance to hate crime

	(1) Low distance	(2) High distance
MCS	-0.569***	-0.281
	(0.165)	(0.196)
PHQ-4	-0.435**	-0.206
	(0.178)	(0.143)
Life satisfaction	-0.109	0.021
	(0.140)	(0.148)
Intention to stay	0.035	0.015
	(0.038)	(0.028)
Number of observations	585	620

Note: Table 3.4 displays the effect of xenophobic attacks on refugees' mental health, disaggregated by geographic distance to the focal hate crime. Column (1) displays point estimates for refugees with low geographic distance to the focal hate crime, while column (2) displays results for refugees with high geographic distance. The coefficients correspond to coefficient estimates of a local linear regression of the mental health outcome on an indicator which is equal to one if a xenophobic attack occurred as well as the temporal relative distance to the attack, allowing for differential trend before and after the xenophobic attack. Throughout, we use triangular kernels and the bandwidth around the cutoff is ± 90 days. The outcomes have been standardized to have mean zero and standard deviation one. The standard errors are clustered on the day relative to the xenophobic attack and are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

3.5.5 Country-specific human capital as a mediator

In the following, we investigate Becker and Rubinstein's (2011) hypothesis that individuals with higher ability are less likely to display large emotional responses to hate crime. Becker and Rubinstein (2011) argue that higher ability allows individuals to align the subjective and objective likelihood of being harmed by hate crime better, i.e., that individuals with higher ability are potentially more likely to obtain and process relevant information necessary to align the objective and subjective likelihood of being harmed by hate crimes. Moreover, for migrants, or more precisely, refugees, we argue that it is country-specific human capital that matters for this process. Thus, we test whether language proficiency and contact to native residents moderates the mental health response to hate crimes.

For this, we distinguish between refugees who speak, read, and write German at least averagely from those who report lower levels of language proficiency. With respect to social capital, we distinguish between refugees who have at least weekly contact with German natives and those who have less contact.[18] Additionally, we distinguish between human capital that is difficult–more costly–to acquire and low-cost human capital. This also allows us to interact the country specific human capital with "general ability" and test for complementarities in the respective relationship. We argue that reading German and having frequent contact with Germans are easier to acquire (low-cost) dimensions of country-specific human capital than writing and speaking German (high-cost).

Table 3.5 and Table 3.6 display the results for the stock of low- and high-cost country-specific human capital, respectively. In Table 3.5, columns (1) and (2) show the results for having at least weekly contact with Germans versus less frequent contact. Columns (3) and (4) display the results for being able to read German at least averagely versus worse than averagely. The results in Table 3.5 suggest a lower effect of hate crime on the mental health of refugees who have frequent contact with Germans and are better able to read German. In effect sizes, the difference between refugees with low versus high country-specific human capital is 5.6 percentage points of a standard deviation for the MCS score. This corresponds to a difference of 15.9% relative to the effect size for those who have less frequent contact with Germans. On the other hand, the difference is 15.1 percentage points for between those who read German at least averagely and those who read German below averagely for the PHQ-4 score. This is equivalent to 42.5% relative to the effect size for those who read less than averagely German. However, the difference between those who read German at

[18]Note that the frequency of having contact with German natives may change as a result of the hate crimes. However, we theorize that this may not be the case regarding the command over the German language.

least averagely and those who read German less than the average is smaller and points in the opposite direction.

Table 3.5: The effect of hate crime on refugees' mental health, conditioning on low cost country-specific human capital

	Contact with Germans		Reading German	
	Yes	No	Yes	No
	(1)	(2)	(3)	(4)
MCS	-0.349**	-0.415***	-0.410***	-0.369**
	(0.169)	(0.152)	(0.135)	(0.164)
PHQ-4	-0.250*	-0.344**	-0.204*	-0.355*
	(0.149)	(0.168)	(0.116)	(0.186)
Number of observations	654	550	622	582

Note: Table 3.5 displays the effect of hate crime on refugees' mental for refugees commanding over low-cost country-specific human capital. We conjecture that "Contact with Germans" and "Reading German" are low-cost country-specific human capital. We distinguish between refugees that command over the country specific human capital ("Yes") or not ("No"). Refugees have contact with Germans if they have contact with Germans on a weekly basis. Refugees command about the skill "Reading German" if they read German at least averagely. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights and a bandwidth of 90 days. The outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

In Table 3.6, columns (1) and (2) display results for individuals who write German at least averagely and less than averagely, while columns (3) and (4) display the results for individuals who speak German at least averagely and less than averagely. Here, the pattern is even more pronounced. The differences uniformly suggest that effect sizes are considerably smaller for refugees with high levels of country-specific human capital. For the PHQ-4 score, results suggest that we can not reject the absence of an effect of hate crime on the refugees' mental health for those who write and speak German at least averagely. Moreover, the difference in effect sizes amounts to 30.8 percentage points of a standard deviation for the PHQ-4 score between those who speak German at least averagely and those who speak German worse than averagely. This corresponds to 69% relative to the effect size for those who speak German worse than averagely. On the other hand, the difference in effect sizes is 14.7 percentage points for the MCS score between those who speak German at least averagely and those who speak German worse than averagely. This is equal to 34.2% relative to the baseline.

Table 3.6: The effect of hate crime on refugees' mental health, conditioning on high cost country-specific human capital

	Writing German		Speaking German	
	Yes	No	Yes	No
	(5)	(6)	(7)	(8)
MCS	-0.277**	-0.449***	-0.283**	-0.430**
	(0.123)	(0.169)	(0.121)	(0.175)
PHQ-4	-0.133	-0.420**	-0.139	-0.447**
	(0.121)	(0.187)	(0.113)	(0.188)
Number of observations	603	601	647	557

Note: Table 3.6 displays the effect of hate crime on refugees' mental health for refugees commanding over high-cost country-specific human capital. We conjecture that "Speaking German" and "Writing German" are high-cost country-specific human capital. We distinguish between refugees that command over the country specific human capital ("Yes") or not ("No"). Refugees command about the skill "Reading German" or "Writing German" if they reply that they speak or read German at least averagely. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights and a bandwidth of 90 days. The outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

Our results indicate clearly that refugees, who have more country-specific human capital experience lower mental health responses to hate crime. Further, this is complementary to the costs of acquiring this country-specific human capital. This is consistent with Becker and Rubinstein (2011) and clearly suggests a high relevance for the ability to acquire information about the respective hate crimes.

3.6 Conclusion

Considering that both hate crimes and the number of refugees are strongly increasing, it is of utmost importance to estimate the costs associated with the hate crimes targeting refugees. Therefore, this chapter shows that hate crimes have a strong and negative effects on refugees' mental health, as measured by the MCS and PHQ-4 score.

The effects are sizable in magnitude and stronger for refugees living in close geographic proximity to the focal hate crime. Our results suggest that mental health shocks exist only temporarily. However, since the feeling of increased uncertainty and insecurity leads to worse economic choices (Becker and Rubinstein, 2011; Schilbach et al., 2016) and because these shocks potentially extend to the next generation (Almond and Currie, 2011; Almond et al., 2018), we argue that these shocks may have important long-run consequences.

We also characterize Becker and Rubinstein's (2011) conjecture that it is high ability individuals who are less prone to hate crimes by showing how important it is to acquire information about these incidences. This "information channel" interacts with the opportunity costs of acquiring this ability adding to the evidence on the importance of country-specific human capital for migrants and refugees. This is particular important for refugees since these are permanent migrants, able to accrue the returns to country-specific human capital over longer time horizons (Dustmann and Glitz, 2011).

In contrast to mental health, we find no effect of hate crime on refugees' life satisfaction or intention to stay forever in Germany. This result stands in clear contrast to the previous literature, which considers the effect of hate crime on economic immigrants' integration (e.g., Deole, 2019; Steinhardt, 2018). The contrasting results may be explained by the inherent differences between refugees and economic migrants, thus reinforcing the importance of distinguishing between these groups.

Our results have very important policy implications. Mental health is a central determinant of individual's well-being, and physical integrity is a basic constitutional right. Further, impaired mental health as a result of perceived hate crimes may have substantial negative effects for refugees that may harm integration in the host country in the long-run. Besides this first-order effect, slow integration of refugees creates substantial negative externalities and fiscal costs for the host societies. As a consequence, our results ask for increased attention towards the mental health needs of refugees being victims of hate crime. In addition, refugees' integration success depends on the host societies' attitudes towards refugees (Ther, 2019) and hate crimes are the most severe form of refusal. If host countries wish to integrate refugees, they should make every effort to create equal opportunities and social cohesion.

Appendix

3.A Additional tables

Table 3.A.1: Attacks against refugee shelters

No.	Date	Place	State	Type of crime	Right-wing
1	01./01/2016	Nienburg/Saale	ST	Insult §185 StGB	x
2	01/01/2016	Merseburg/Saale	ST	Sedition §130 StGB	x
3	01/01/2016	Wernigerode	ST	Property damage §304 StGB	x
4	01/01/2016	Assamstadt	BW	Grievous bodily harm §224 StGB	x
5	01/01/2016	Werbach	BW	Use of symbols of unconstitutional organizations §86a StGB	x
6	01/01/2016	Ruppertshofen	BW	Use of symbols of unconstitutional organizations §86a StGB	
7	01/01/2016	Zeven	NI	Grievous bodily harm §224 StGB	
8	01/01/2016	Leverkusen	NW	Grievous bodily harm §224 StGB	x

Note: This table is based on administrative data on hate crimes against refugee shelters (all entries for January 1, 2016), which is published by the German Federal government on a quarterly basis. Source: BKA data (2016-2018).

Table 3.A.2: Factors of a principal component analysis of the subscales of the SF-12 questionnaire

	PCS score (1)	MCS score (2)
Physical Fitness	0.791	0.084
General Health	0.740	0.281
Bodily Pain	0.831	0.194
Role Physical	0.823	0.313
Mental Health	0.155	0.823
Role Emotional	0.544	0.605
Social Functioning	0.494	0.577
Vitality	0.108	0.700

Note: Table 3.A.2 displays the factor loadings of a principal component analysis of the subscales of the SF-12 questionnaire. The factor analysis has been performed on all first time respondents of the IAB-BAMF-SOEP Refugee Survey in 2016 and 2017. Column (1) displays the corresponding factor loadings for the first factor, which corresponds to the PCS score. Column (2) displays the factor loadings of the second factor, which corresponds to the MCS score.

Table 3.A.3: Factor loadings on the first factor of a principal component analysis of the items of the PHQ-4 inventory

	PHQ-4 score
	(1)
Little Interest	0.598
Melancholy	0.845
Anxiety	0.844
Worrying	0.786

Note: Table 3.A.3 displays the factor loadings of a principal component analysis on the items of the PHQ-4 inventory. scale under consideration. Column (1) displays the corresponding factor loadings for the first factor. The factor analysis has been performed on the first time respondents of the IAB-SOEP-BAMF Refugee Survey in 2016.

Table 3.A.4: The effect of hate crime on refugees' mental health, life satisfaction and intention to stay, controlling for predetermined characteristics

	(1)	(2)	(3)
MCS	-0.337***	-0.361***	-0.329***
	(0.118)	(0.119)	(0.119)
PHQ-4	-0.291***	-0.284***	-0.290***
	(0.107)	(0.102)	(0.105)
Life satisfaction	-0.030	-0.031	-0.023
	(0.097)	(0.103)	(0.097)
Intention to stay	0.029	0.023	0.030
	(0.022)	(0.023)	(0.022)
Refugees' predetermined characteristics	✓		✓
Regional predetermined characteristics		✓	✓
Number of observations	1215	1215	1215

Note: Table 3.A.4 displays the effect of hate crime on refugees' mental health, life satisfaction and intention to stay, controlling for individual or county level characteristics. Columns (1), (2), (3), and (4) display point estimates and corresponding standard errors for estimations including predetermined individual, regional as well as individual and regional characteristics, respectively. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators in addition to the respective predetermined characteristics. We use triangular weights and a bandwidth of 90 days. Non-binary outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

Table 3.A.5: The effect of xenophobic attacks on refugees' mental health, accounting for multiply treated

	(1) 14 days	(2) 30 days	(3) 45 days	(4) 60 days	(5) 90 days
MCS	-0.280***	-0.369***	-0.422***	-0.636***	-0.713***
	(0.099)	(0.118)	(0.143)	(0.131)	(0.145)
PHQ-4	-0.268***	-0.285***	-0.331***	-0.483***	-0.550***
	(0.100)	(0.104)	(0.125)	(0.110)	(0.121)
Life satisfaction	-0.153*	-0.040	-0.081	-0.063	-0.078
	(0.086)	(0.102)	(0.127)	(0.129)	(0.142)
Intention to stay	0.011	0.023	0.037	0.041	0.030
	(0.023)	(0.022)	(0.028)	(0.033)	(0.026)
Number of observations	1333	1215	1098	982	770

Note: Table 3.A.5 displays the effect of xenophobic attacks on refugees' mental health. We argue that estimates may be downward biased if gate crimes are clustered in time. Therefore, we drop observations who experienced a second hate crime shortly before the focal hate crime for different time periods. Column (1) drops observations who experience a second hate crime in a fourteen day period preceding the focal hate crime. Columns (2), (3), (4), and (5) display the results for a thirty (baseline estimation), forty-five, sixty, and ninety day period, respectively. The coefficients correspond to coefficient estimates of a local linear regression of the mental health outcome on an indicator which is equal to one if a xenophobic attack occurred as well as the temporal relative distance to the attack, allowing for differential trend before and after the xenophobic attack. Throughout, we use triangular kernels and the bandwidth around the cutoff is 90 days. The outcomes have been standardized to have mean zero and standard deviation one. The standard errors are clustered on the day relative to the xenophobic attack and are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

Table 3.A.6: The effect of hate crime on refugees' mental health, allowing for a quadratic trend in the running variable

	MCS	PHQ-4	LS	ITS
	(1)	(2)	(3)	(4)
Effect of hate crime	-0.374**	-0.339***	0.080	0.028
	(0.155)	(0.129)	(0.151)	(0.031)
Number of observations	1393	1410	1244	1371
MSE-optimal bandwidth	114.445	117.850	92.039	109.783

Note: Table 3.A.6 displays the effect of hate crime on refugees' mental health, life satisfaction and intention to stay allowing for a quadratic trend. Columns (1), (2), (3), and (4) display point estimates and corresponding standard errors for the MCS score, the PHQ-4 score as well as life satisfaction and refugees' intention to stay, respectively. The coefficient estimates are based on a local polynomial regression, in which we regress the respective outcome on an indicator for a hate crime, a quadratic trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights. Non-binary outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the running variable level, distance in days to the focal hate crime, and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

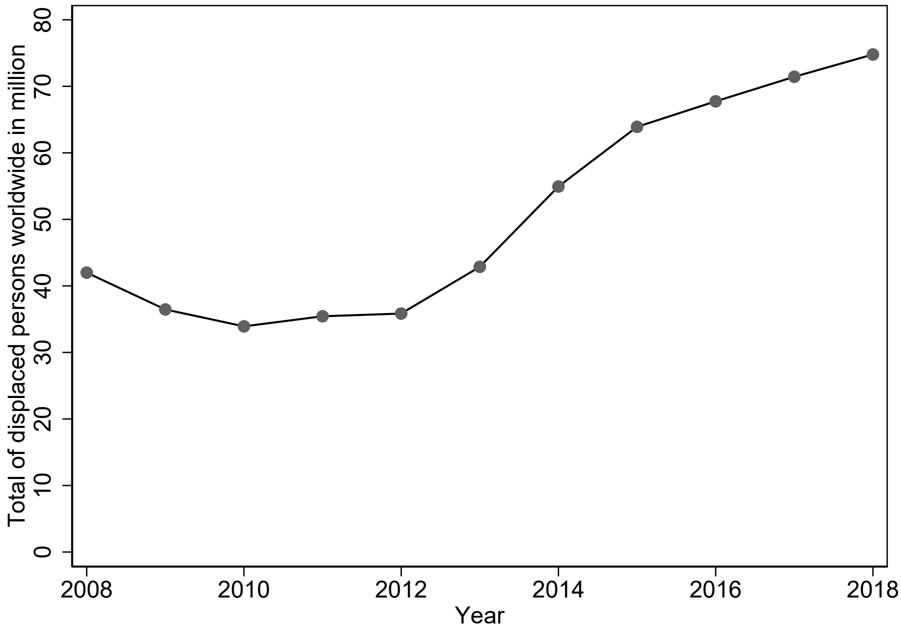
Table 3.A.7: Clustering on the county level

	MCS	PHQ-4	LS	ITS
	(1)	(2)	(3)	(4)
Effect of hate crime	-0.368*	-0.284**	-0.040	0.023
	(0.192)	(0.130)	(0.116)	(0.023)
Number of observations	1215	1215	1215	1215

Note: Table 3.A.7 displays the effect of hate crime on refugees' mental health, life satisfaction and intention to stay, clustering the standard errors on the level of the county. Columns (1), (2), (3), and (4) display point estimates and corresponding standard errors for the MCS score, the PHQ-4 score as well as life satisfaction and refugees' intention to stay, respectively. The coefficient estimates are based on a local linear regression, in which we regress the respective outcome on an indicator for a hate crime, a linear trend in the running variable, which is allowed to differ before and after the focal hate crime, and day of week as well as month of year indicators. We use triangular weights and a bandwidth of 90 days. Non-binary outcomes are standardized to have mean zero and a standard deviation of one. Standard errors are clustered on the county level and are displayed in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Source: SOEP, v34.

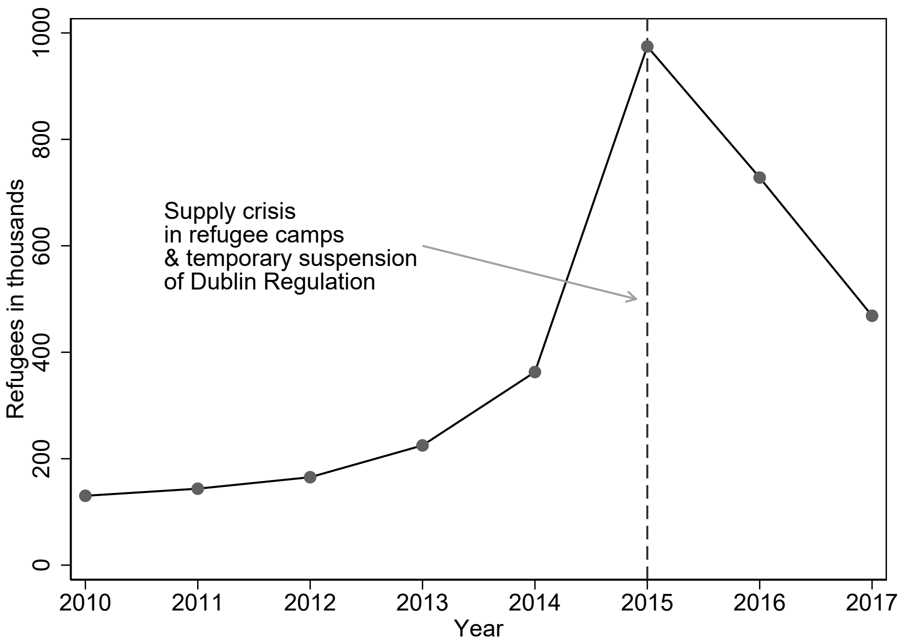
3.B Additional figures

Figure 3.B.1: Time trend in the number of displaced persons



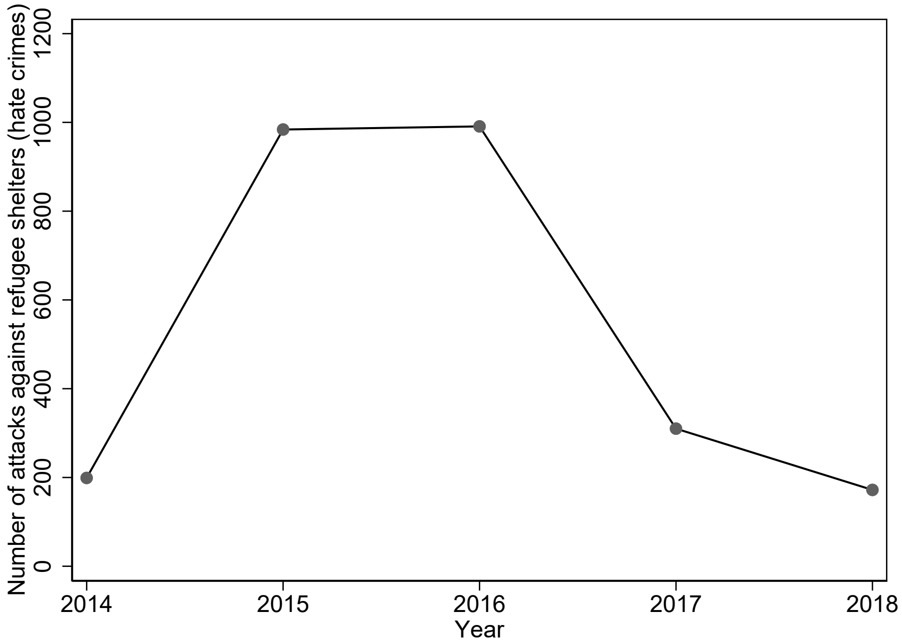
Note: Figure 3.B.1 plots the number of displaced persons worldwide from 2008 to 2018. Source: UNHCR (2009) to UNHCR (2019).

Figure 3.B.2: Number of asylum seekers in Germany



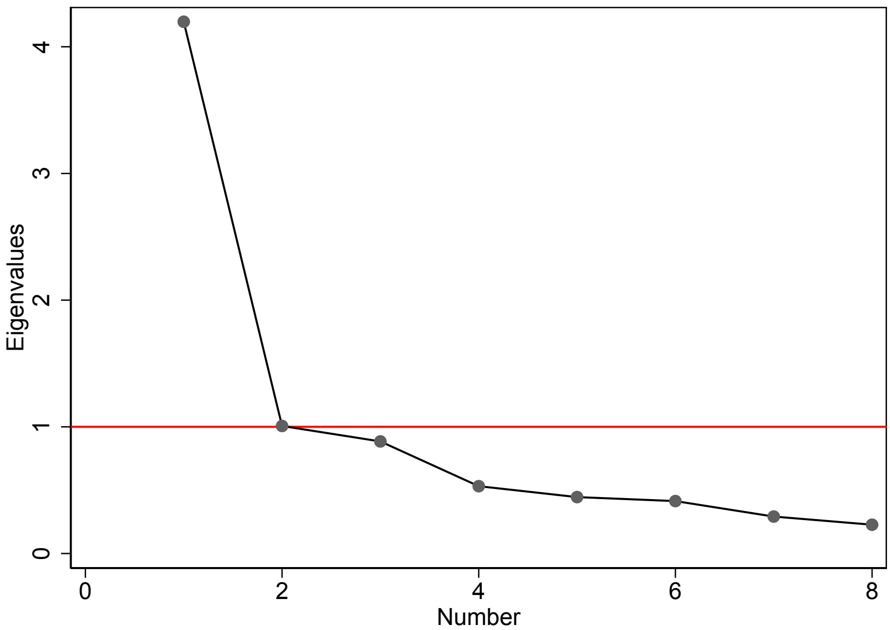
Note: Figure 3.B.2 plots the number asylum seekers from 2010 to 2017. Source: Federal Statistical Office of Germany (2019).

Figure 3.B.3: Number of attacks against refugee shelters over time



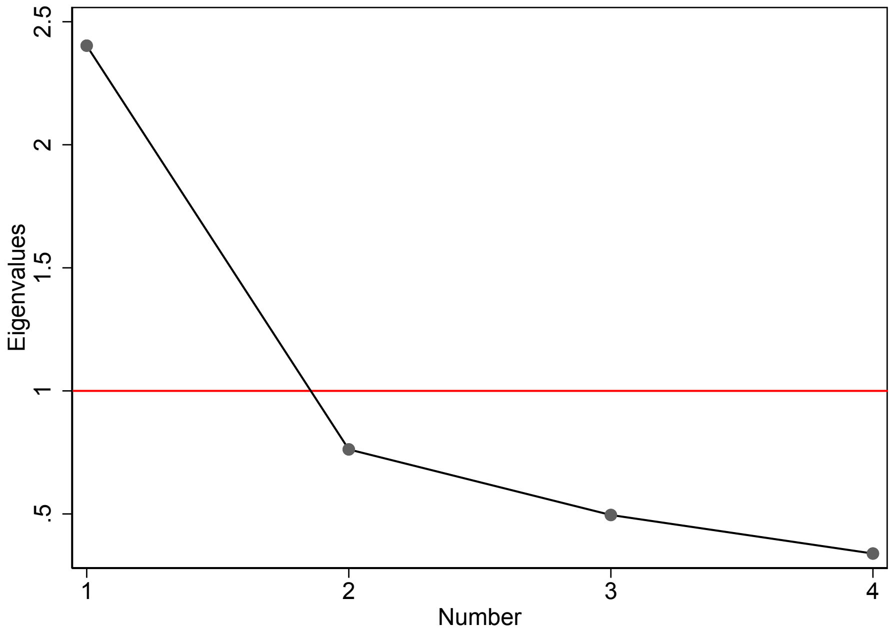
Note: Figure 3.B.3 plots the number of attacks against refugee shelters. Source: Bundestag (2016).

Figure 3.B.4: Scree plot for principal component analysis of the subscales of the SF-12 questionnaire



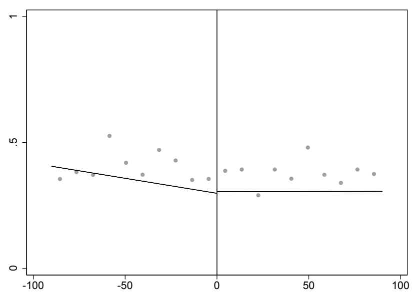
Note: Figure 3.B.4 plots the factors and the corresponding Eigenvalues after a principal component analysis of the subscales of the SF-12 questionnaire. The horizontal red line corresponds to Eigenvalues of one. Source: SOEP, v34.

Figure 3.B.5: Scree plot for principal component analysis of the items of the PHQ-4 inventory

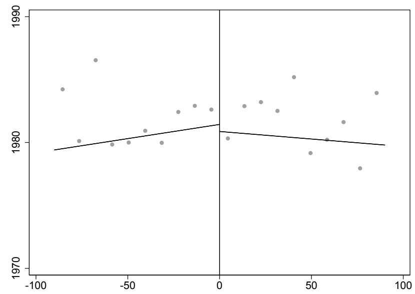


Note: Figure 3.B.5 plots the factors and the corresponding Eigenvalues after a principal component analysis of items of the PHQ-4 inventory. The horizontal red line corresponds to Eigenvalues of one. Source: SOEP, v34.

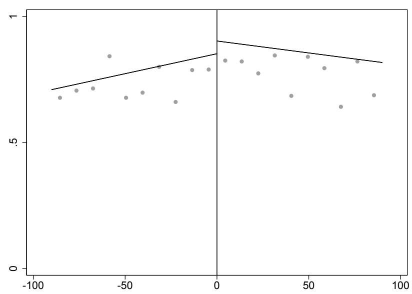
Figure 3.B.6: Test of continuity assumption



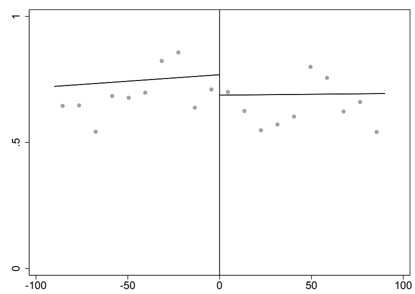
(a) Female



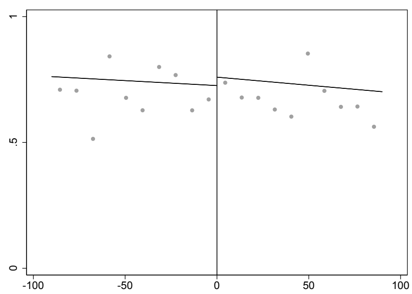
(b) Year of birth



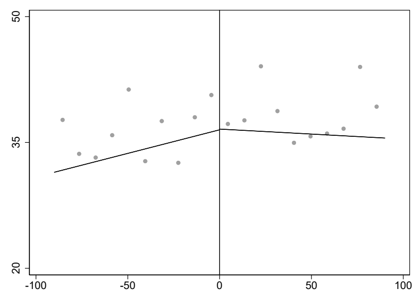
(c) Country of origin SIA



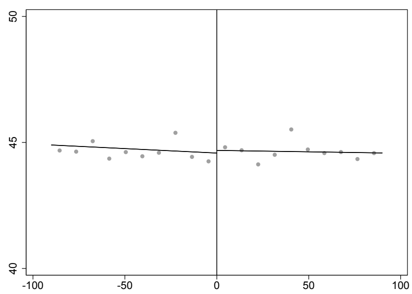
(d) Child



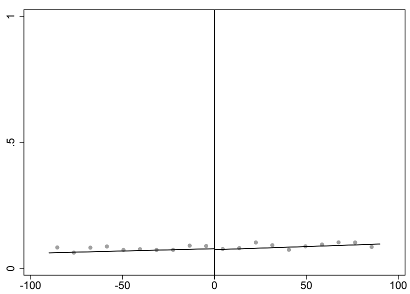
(e) Married



(f) BIP per capita in county



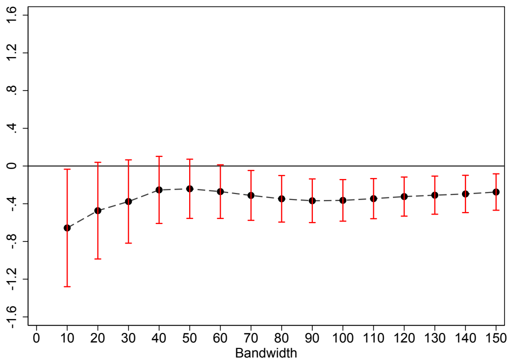
(g) Average age in county



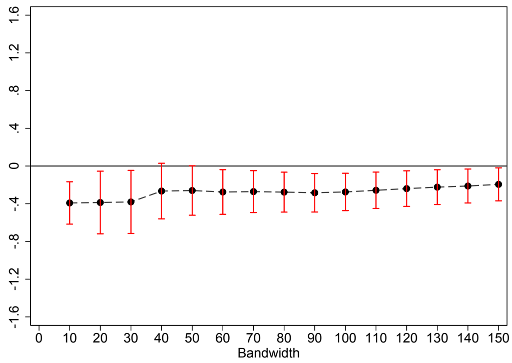
(h) Share foreign in county

Note: Figures 3.B.6a to 3.B.6h display visual results for the test for continuity of predetermined characteristics around the focal hate crime. Throughout, the bandwidth is chosen to be 90 days. The dots correspond to a binned scatterplots. The vertical bars are 95% confidence intervals for the means of the bins, based on standard errors that are clustered on the running variable. The linear fit corresponds to a local linear regression with a triangular kernel as in Equation 3.1.

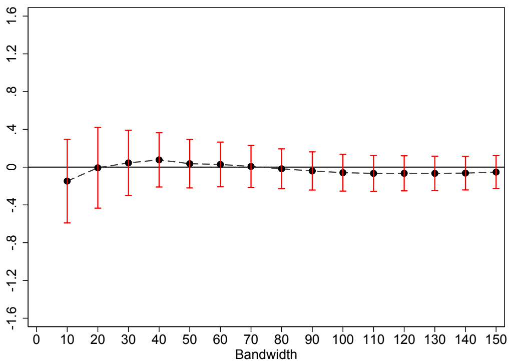
Figure 3.B.7: Sensitivity of the estimates to the bandwidth choice



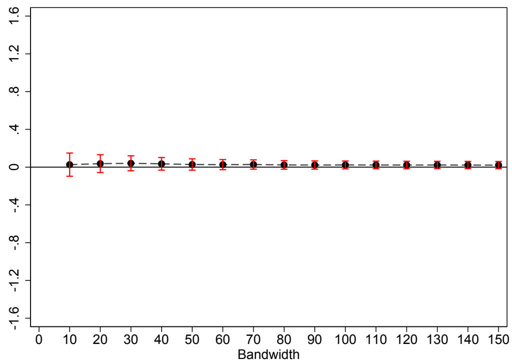
(a) MCS



(b) PHQ-4



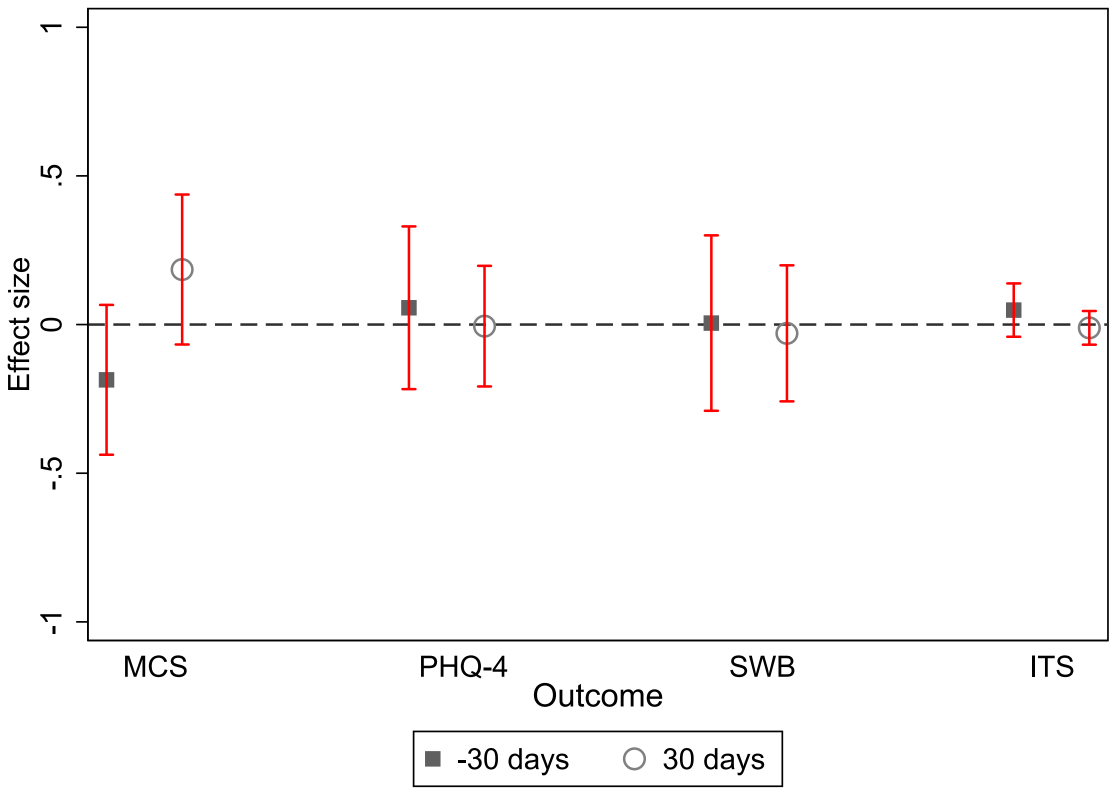
(c) Life satisfaction



(d) ITS

Note: Figures 3.B.7a to 3.B.7d display the effect of xenophobic attacks on the MCS and PHQ-4 score, life satisfaction as well as intention to stay conditional on the bandwidth choice, respectively. In each figure, a dot corresponds to a point estimate corresponding to bandwidth choice each. The estimates correspond stem from a local linear regression of the respective mental health outcome on an indicator for xenophobic attacks and a linear trend in the running variable, which is allowed to vary before and after the cutoff. We used triangular kernels. The red bars display 95% confidence bands. Throughout, we clustered standard errors on the relative distance to the xenophobic attack. Source: SOEP, v34.

Figure 3.B.8: Placebo estimates



Note: Figure 3.B.8 displays the point estimates and 95% confidence intervals of placebo tests. For each mental health outcome, the left estimates correspond to point estimates of placebo regressions, pretending the xenophobic attack happened 30 days before the actual xenophobic attack. The right estimates display the respective estimates pretending the xenophobic attack happened 30 days after the actual xenophobic attack. Source: SOEP, v34.

CHAPTER 4

COVID-19: A crisis of the female self-employed*

We investigate how the economic consequences of the pandemic and the government-mandated measures to contain its spread affect the self-employed – particularly women – in Germany. For our analysis, we use representative, real-time survey data in which respondents were asked about their situation during the COVID-19 pandemic. Our findings indicate that among the self-employed, who generally face a higher likelihood of income losses due to COVID-19 than employees, women are 35% more likely to experience income losses than their male counterparts. We do not find a comparable gender gap among employees. Our results further suggest that the gender gap among the self-employed is largely explained by the fact that women disproportionately work in industries that are more severely affected by the COVID-19 pandemic. Our analysis of potential mechanisms reveals that women are significantly more likely to be impacted by government-imposed restrictions, e.g., the regulation of opening hours. We conclude that future policy measures intending to mitigate the consequences of such shocks should account for this considerable variation in economic hardship.

*This chapter is joint work with Alexander Kritikos and Johannes Seebauer. A previous version was published as SOEPpaper on Multidisciplinary Panel Data Research 1108 and Discussion Papers of DIW Berlin 1903 in 2020. This version has been published as CEPA Discussion Paper 27 and GLO Discussion Paper 788 in 2021. This chapter is forthcoming in the *Journal of Population Economics*.

4.1 Introduction

The unprecedented shutdown of businesses in specific industries, social distancing guidelines, and overall insecurity caused by the COVID-19 pandemic resulted in the temporary halt of major parts of the economy in many countries in 2020, with dire consequences for these economies (Milani, 2021). The service sector, which often necessitates physical proximity, was particularly affected (Barbieri et al., 2020). At the same time, this sector depends more on self-employed individuals than the manufacturing sector, where the vast majority of workers are employees. In particular, self-employed women are more likely to work in service industries than self-employed men: According to the OECD (2017), 91% of self-employed women and 68% of self-employed men in Germany worked in the service sector in 2016.

The COVID-19 pandemic initiated a public debate as to what extent the female working population experienced greater income and employment reductions. This is particularly relevant since women are often the primary caretakers in the family and, as such, were also confronted with the closure of schools and daycare centers (Alon et al., 2020). However, the debate revolving around the gender gap and the impact of the COVID-19 pandemic does not, thus far, differentiate between different employment forms, although initial descriptive evidence points to stronger negative effects for self-employed women (see e.g. Ifo Institute and forsa (2020) for Germany and Kalenkoski and Pabilonia (2020) for the U.S.). In this paper, we investigate whether women in self-employment and employment are more severely affected by the economic consequences of the COVID-19 pandemic and associated non-pharmaceutical interventions (NPI) than men. To the best of our knowledge, we are the first to explicitly contrast the experience of the self-employed with employees during the COVID-19 pandemic and, by doing so, to identify where gender disparities occurred as a consequence of the pandemic.

The particular focus on self-employed individuals is warranted by the increasing relevance of self-employment and entrepreneurship for modern economies. For example, in Germany, around 4.2 million individuals – about ten percent of the working population – are self-employed, running diverse businesses either without or with further employees, often micro-businesses with up to 10 employees. In sum, the self-employed contribute substantially to the economic development of the country (Audretsch et al., 2020). It is further important to note that, while there is still a significant gender gap among the self-employed, the share of women has been increasing steadily since the turn of the century (Fritsch et al., 2015).

Our study proceeds in three steps. First, we contextualize our analysis on the comparison between female and male workers in both employment forms by investigating the differential impact of the COVID-19 pandemic on the self-employed and employees. Second, in our main analysis, we examine the gender gap in the effect of the pandemic on labor market outcomes, thereby focusing on the self-employed. Third, we provide evidence for potential mechanisms driving the observed gender differences among the self-employed. For our analysis, we use the Socio-Economic Panel-CoV (SOEP-CoV), a novel data set sufficiently rich to allow for such a comparison, as it enables us to control for individual-level heterogeneity to a large extent. SOEP-CoV surveyed a randomly selected subset of respondents from the SOEP who were asked to answer a wide array of questions about their economic situation, family situation, health, the use of public support instruments, as well as attitudes during the early stages of the COVID-19 pandemic. The SOEP is a representative household panel in Germany that surveys respondents annually since 1984 (Goebel et al., 2019). By design, the SOEP-CoV enables us to link individual respondents to their pre-crisis information. Thus, we can exploit rich information on the respondents, including their pre-crisis household income, education, household characteristics, personality traits, and employment experience, among others. Therefore, we are able to analyze whether individual characteristics that are known to be important determinants of self-employment, influenced outcomes during the COVID-19 pandemic (see e.g. Parker, 2018).

With this data at hand, we perform multivariate analyses, first comparing the gap in labor market outcomes between employed and self-employed respondents. We show that there are significant differences in the influence of the COVID-19 pandemic and associated NPIs on the two employment forms: The self-employed are about 42 percentage points more likely to report losses of gross income than employees and 30 percentage points more likely to report a reduction in working hours. Turning to gender differences in the influence of the COVID-19 pandemic, we find that self-employed women are about one-third more likely to experience income losses due to the

COVID-19 pandemic compared to self-employed men. We do not find a comparable gender gap among employees.

We then decompose the gender gap in the probability of income losses among the self-employed using the Gelbach decomposition (Gelbach, 2016), thus allowing us to decompose different sets of covariates into their individual contribution to the gender gap. We show that the gender gaps in the probability of income losses and reductions in working hours due to the COVID-19 pandemic are driven by the fact that self-employed women are disproportionately active in industries that are more severely affected by the COVID-19 pandemic. We do not find such evidence for employees.

Lastly, we provide evidence for a channel driving the gender gap among the self-employed. We find that self-employed women are 20 percentage points more likely to be affected by regulations due to the COVID-19 pandemic.

We show that our results are, once again, driven by the disproportionate sorting of self-employed women into industries that were more severely restricted by the NPIs implemented. Moreover, we present evidence that these restrictions mediate the relationship between industry-sorting and income losses. We also find suggestive evidence that gendered household production contributes to the gender gap in income losses. However, this effect is of second order compared to the contribution of industry affiliation.

We contribute to the literature in several ways: First, we contrast the gender gap between employees and self-employed individuals in the labor market during the early onset of the COVID-19 pandemic. In contrast to related studies relying on the U.S. Current Population Survey (Fairlie, 2020; Kalenkoski and Pabilonia, 2020) or the Canadian Labour Force Survey (Beland et al., 2020), the SOEP-CoV contains information on earnings losses due to the COVID-19 pandemic. Adams-Prassl et al. (2020), who collected their own data, is a notable exception in that they do have information on earnings losses. The authors do not find gender differences in realized job or earnings losses for Germany. While they provide important initial evidence, they do not distinguish between self-employed individuals and employees with respect to the gender gap. This is an important distinction since the labor market in Germany is characterized by stronger rigidities than other countries, limiting the extent to which firms can cut the wages of their employees (e.g. Burda, 2016). Furthermore, policy measures taken by the federal government were mostly aimed at stabilizing the earnings and employment trajectories of employees. By contrast, self-employed individuals, as residual claimants, are more vulnerable to economic shocks like the COVID-19 pandemic.

Second, we contribute to the broader literature on gender gaps in labor markets (e.g. Blau and Kahn, 2017; Goldin et al., 2017; Meara et al., 2020) that documents earnings gaps, which the authors, among others, attribute to selection of women into occupation or sectors that are associated with lower average wages. We complement this literature with our finding that the disproportionate representation of women in certain industries also translates into a gender gap in the impact of the COVID-19 pandemic. Third, our finding that government-mandated regulations are an important driver of the gender gap in the impact of the pandemic on the self-employed constitutes novel evidence in the literature.

Lastly, we also contribute to a strand of literature studying the consequences of the spread of communicable diseases on economic well-being (e.g. Karlsson et al., 2014; Barro et al., 2020; Correia et al., 2020; Velde, 2020). These studies mainly investigate the impact of the 1918 Spanish flu. While providing important insights, these are restricted by limited data due to the historic nature of the event. In this context, our finding that NPIs have unintended consequences for gender equality implies that this variation in economic suffering needs to be accounted for when addressing the ongoing COVID-19 pandemic or any future public health crisis involving communicable diseases of a similar or even greater magnitude.

4.2 Background: The COVID-19 pandemic, policy measures, and female self-employment

In this section, we provide a short summary of policy measures enacted in Germany in the early months of the pandemic, before we relate our study to contemporaneous research on the impact of COVID-19 on self-employment, as well as on the gender gap in self-employment.

4.2.1 Policy measures in the wake of the COVID-19 pandemic

In order to contain SARS-CoV-2, the German government imposed strong restrictions beginning in March 2020, shortly before our period of observation. These NPIs included the closure of schools, daycare centers, restaurants, service companies in the field of personal hygiene, and most shops – with exceptions for grocery stores. All public events were canceled and travel restricted. Meetings in public were limited to two individuals, while people were required to keep a minimum distance of 1.5 meters from other people in public spaces (Federal Ministry of Health, 2020). While these measures were certainly sensible from an epidemiological perspective (e.g. Qiu et al., 2020; Bonacini et al., 2021), more than half of the self-employed experienced sales and income losses in spring 2020 (Kritikos et al., 2020).

The German government introduced several economic policy measures to mitigate the economic consequences of the COVID-19 pandemic. The most prominent policy measure being the expansion of "*Kurzarbeit*", the established short-time work compensation scheme where the employment agency covers up to 67% of employees' net income.[1] As the self-employed are not covered by this instrument, the federal government released an emergency aid package of up to €50 billion for the self-employed. This program supported the self-employed facing strong losses in revenues with lump sum payments of up to €15,000. The use of this payment was limited to covering fixed operating costs and temporarily increased the subjective survival probability (Block

[1] Under this scheme, employers send their employees into short-time work where the Federal Employment Office subsidizes a large portion of the wage costs pertaining to those contractual working hours that employees are not working. This instrument allows employers to keep their workforce through the crisis while protecting employees from losing their jobs, and from major wage losses, see also Cahuc (2019).

et al., 2020). In addition, the self-employed received easier access to unemployment benefits "*Arbeitslosengeld 2*" (Federal Ministry for Economic Affairs and Energy, 2020).

4.2.2 Related research on self-employment

Crisis-related research on self-employment has received considerable attention (see e.g. Doern et al., 2019). On the one hand, a large part of this literature focuses on the question of how individuals decide about venturing new businesses in reaction to a crisis (see e.g. Siemer, 2014) and, on the other hand, the crisis management of existing businesses (see e.g. Davidsson and Gordon, 2016). Much less is known about the magnitude of the impact of crises on the self-employed; existing research is often based on qualitative interviews with retrospective questions (see e.g. Doern, 2016).

In contrast to other crises, the COVID-19 pandemic affects nearly the entire self-employed population, as is documented in contemporaneous research, all of which shows that self-employed individuals suffered significantly from the consequences of the COVID-19 pandemic.[2] For the U.K., Blundell and Machin (2020) show that three out of four self-employed individuals report a reduced work load. While they provide important evidence on the impact of the COVID-19 pandemic on self-employed individuals, they do not consider gender differences in their analysis. Fairlie (2020) documents that the activity of business owners in the U.S. plummeted by 3.3 million, or 22%, during the early stages of the COVID-19 pandemic. Fairlie (2020) also documents considerable race and gender differences in the effects of the COVID-19 pandemic on the number of active small businesses. In contrast to our study, Fairlie (2020) does not have information on income losses. Kalenkoski and Pabilonia (2020), who focus on unincorporated self-employed in the U.S., find that self-employed individuals are about 57 percentage points less likely to be employed in April 2020, compared to February. The authors, like Fairlie (2020), also do not have information on income. Kalenkoski and Pabilonia (2020) likewise document gender differences in the effects of the COVID-19 pandemic on self-employed individuals. Lastly, Beland et al. (2020) report an activity decline of 14.8% for incorporated and 10.1% for unincorporated entities in Canada. They also find gender differences in the impact on COVID-19 on

[2] There are also various studies investigating the effects of the COVID-19 pandemic on overall employment (Forsythe et al., 2020; Chetty et al., 2020; Cajner et al., 2020; Juranek et al., 2020; Coibion et al., 2020; Adams-Prassl et al., 2020).

employment and hours, yet do not analyze this differential impact, nor do they have information on income.

In summary, we expand the analysis on gender differences in the effect of the COVID-19 pandemic on self-employed individuals in two important ways: First, we have information on income losses, in addition to income information from 2019. Second, we provide important evidence that it is the sorting of women into industries that are more strongly affected by the pandemic and associated NPIs that drives the observed gender differences among the self-employed.

Lastly, our study also relates to the literature on gender gaps in self-employment. In most countries, fewer women than men are self-employed (Elam et al., 2019). While the female share of self-employment was as low as 25% at the turn of the century in Germany (Fritsch et al., 2015), it has continuously increased to nearly 35% in 2017 (Günther and Marder-Puch, 2019). This development was also aided by the active promotion of self-employment via start-up subsidies (see e.g. Caliendo and Künn, 2015). The literature documents a variety of reasons for the still existing gender gap in self-employment, ranging from differences in the intergenerational transfer of human capital (see e.g. Georgellis and Wall, 2005), differing influences of age (see e.g. Leoni and Falk, 2010), differing risk attitudes (Caliendo et al., 2014), self-confidence (see e.g. Koellinger et al., 2013), or the willingness to compete (see e.g. Bönte and Piegeler, 2013), while there is also substantial heterogeneity in employment decisions both among women and between women and men (Patrick et al., 2016). Certainly, these differences may inform the implications of our findings for the development of female self-employment post-pandemic.

4.3 Data

In this section we briefly describe our data set and discuss the outcome variables used in the analysis. We then provide descriptive statistics of these outcome variables.

4.3.1 SOEP-CoV

For our analysis, we use a unique data source to estimate the effect of the COVID-19 pandemic on the self-employed. The SOEP-CoV survey was launched in April

2020 to investigate the socio-economic consequences of the COVID-19 pandemic in Germany. In the first part of this special survey, respondents, interviewed in nine waves between April and July 2020, were asked about their economic status, family situation, health information, and attitudes during the COVID-19 pandemic (Kühne et al., 2020). Importantly, the SOEP-CoV questionnaire includes a set of questions targeting self-employed individuals.

What makes the SOEP-CoV particularly useful is its integration into SOEP. The Socio-Economic Panel (SOEP) is a representative, longitudinal survey of households in Germany that started 1984 and is administered to households and the households' members on a yearly basis since then.[3] As of 2020, the SOEP includes approximately 20,000 households with more than 30,000 adult household members. The SOEP contains information on the households and its members' economic situation, education, and attitudes, among others (Goebel et al., 2019).

The respondents surveyed in the SOEP-CoV are a random subset of the SOEP population. Thus, it combines the wealth of longitudinal, pre-pandemic information from the SOEP with a wide array of questions that are related specifically to the COVID-19 pandemic. These unique features make the SOEP-CoV the ideal data set to analyze our research questions. For our analysis, we focus on individuals who are either gainfully employed (part- and full-time) or self-employed. We do not consider self-employed individuals who identified as helping family members in 2019. The distribution of observations of our final sample over calendar weeks in 2020 is displayed in Figure 4.B.1 in the appendix.

4.3.2 Outcome variables

In our analysis, we investigate the differential influence of the COVID-19 pandemic by self-employment status and gender. We focus on the likelihood of experiencing a decrease in income (gross earnings), working hours, and working at least partially from home due to the COVID-19 pandemic. In addition, we also have information on the magnitude of losses of monthly income and reductions in weekly working hours. These outcomes jointly determine how individuals have experienced the COVID-19 crisis to

[3]We use the SOEPv35. DOI: 10.5684/soep-core.v35. In addition, we use the preliminary data of the SOEP for 2019.

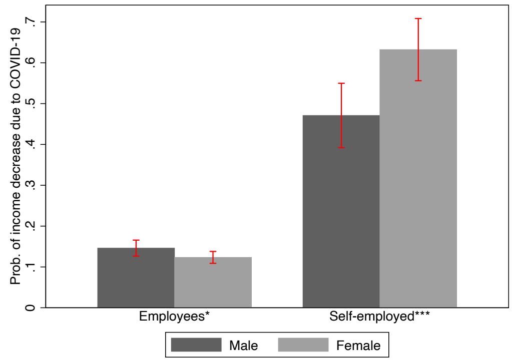
a significant degree and allow for examining differences between employees and the self-employed. Importantly, the questions on income losses, reductions in working hours, and remote work are framed causally. That is, respondents are explicitly asked whether, and to what extent, income and hours worked have changed due to the pandemic. Similarly, they are asked whether they are working from home due to the pandemic, either in part or completely.

While employees are partially protected from income losses in the short-run, when they have fixed employment contracts, this does not apply to the self-employed. The main mechanisms through which employees can face changes in income and working hours are job losses and participation of their employer in short-time work schemes. Furthermore, employees and self-employed individuals may select into different industries. To the extent that these industries are hit by the crisis to varying degrees, the likelihood of reductions in incomes and working hours will differ. The same argument applies to gender differences. To the extent that women select into different industries and occupations than men, along with the extent that these are differently affected by the pandemic, its effect on income and hours will be different. Finally, the potential for working remotely vastly differs across sectors and jobs (Gaudecker et al., 2020; Alipour et al., 2020; Dingel and Neiman, 2020). While front-line workers continued to be potentially exposed to the virus throughout the pandemic, if production was not completely stopped, it was more easily possible for individuals in office jobs to do their work partly, if not completely, from home. By contrast, the arts and entertainment industry, where remote work is nearly non-existent, came to an almost complete halt. Thus, in our main analysis, we shed light on the heterogeneous influence of the COVID-19 pandemic on these core outcomes, which shape the experience of the workforce during the COVID-19 pandemic. Other variables used in the analysis are described in Table 4.A.1 in the appendix.

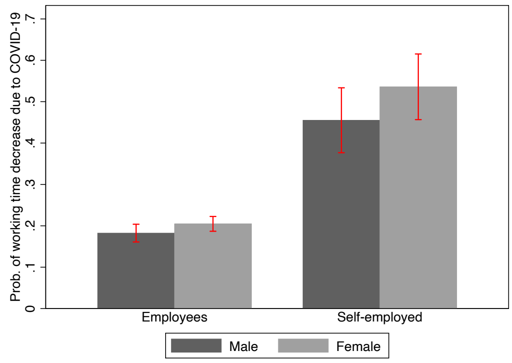
4.3.3 Descriptive statistics on outcomes at the extensive margin

Tables 4.A.2 to 4.A.4 in the appendix show summary statistics for our analysis sample. The sample is restricted to those individuals for whom the full set of control variables used is available. Importantly, they describe how self-employed individuals were af-

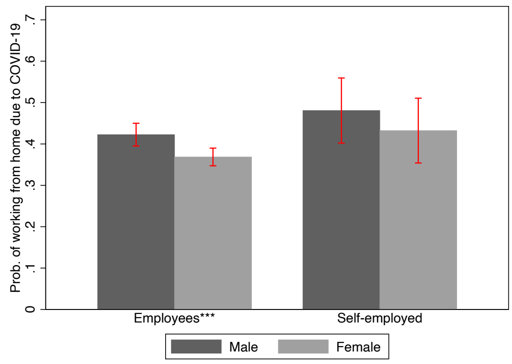
Figure 4.1: Gender comparison of raw differences in probabilities of labor market outcomes



(a) Probability of income decrease



(b) Probability of working time decrease



(c) Probability of remote work

Note: Figures 4.1a to 4.1c display the raw differences in the probability of labor market outcomes over employment status and gender, respectively. Vertical bars correspond to 95% confidence intervals. The stars next to the respective employment group indicate whether the mean differences by gender within the groups are statistically significant and read * p<0.10, **p<0.05, *** p<0.01. Details are displayed in Table 4.A.2 to Table 4.A.4 for details.

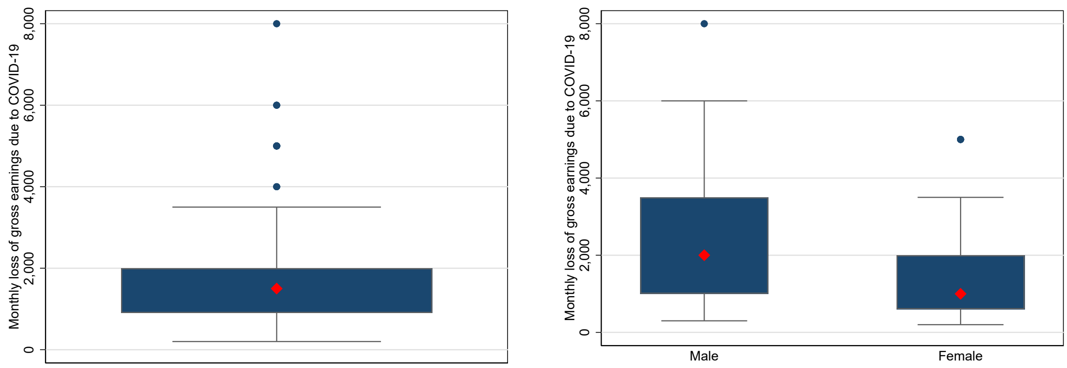
fected by the pandemic in comparison to employees with respect to our outcomes of interest, and how these experiences differ by gender in both employment forms. Figures 4.1a to 4.1c illustrate this difference. The probability of facing reductions in income and working hours is considerably larger among the self-employed than among employees. Around 55% of self-employed individuals report a decline in income and around 50% in working hours, while this is the case for only 13% of employees with respect to income and 20% of them with respect to working hours. A drop in demand directly affects the income and workload of self-employed individuals, whereas income and working hours of employees are affected by a sales decrease in their firms only if they are sent into short-time work or laid off. While job losses following the initial COVID-19 pandemic lockdown are rare in Germany, at least when compared to the experience of other countries (Adams-Prassl et al., 2020), the instrument of short-time work is used extensively.[4] Although the difference is notably smaller, remote work as a direct consequence of the pandemic is also more common among the self-employed (with 46%) than among employees (39%).

Figure 4.1 also shows striking patterns of gender differences in the outcome variables. Most notably, there is a significant gender gap within the group of self-employed individuals: 63% of self-employed women faced income losses as opposed to 47% of their male counterparts. At the same time, 54% of self-employed women and 46% of self-employed men reduced their working hours. With respect to remote work, the gender gap is smaller and, in fact, inverts with men being more likely to work from home than women.

These gender gaps, however, are not replicated among employees. Here, the gender difference in the probability of income losses amounts to roughly two percentage points and inverts. The gender gap in the probability of working from home is similar in magnitude to that of the self-employed. Thus, there is a significant self-employment gap in the outcomes of interest with sizeable gender differences that are concentrated among the self-employed.

[4]See Table 4.A.9 and Section 4.4.1 for a discussion of job loss due to the pandemic.

Figure 4.2: The distributions of absolute monthly losses in gross earnings among self-employed individuals



(a) All (b) Gender differences

Note: Figures 4.2a and 4.2b display boxplots for monthly income losses among all self-employed individuals as well as self-employed men and women. The red marker indicates the median. The upper and lower end of the box display the range between the 25th and 75th percentiles. The whiskers span all data points within 1.5 inter-quartile range of the nearer quartile. Blue dots indicate observations outside the whiskers.

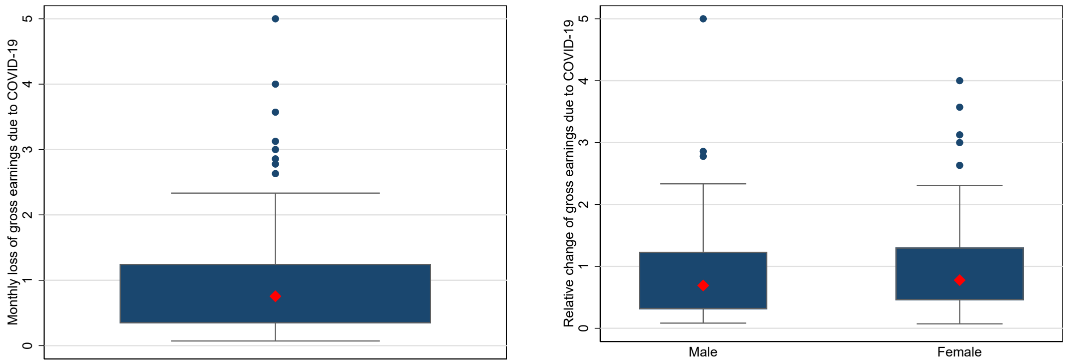
4.3.4 Descriptive statistics on decreases in income and hours at the intensive margin

We also provide descriptive evidence on the magnitude of decreases in income and working hours among the self-employed, beginning with the magnitude of losses in monthly earnings.[5] Figure 4.2 displays the boxplots for monthly absolute income losses for all self-employed individuals as well as separately for women and men. The median and mean of monthly income losses due to the COVID-19 pandemic are €1,500 and €3,020.67 for all self-employed individuals, respectively. Self-employed men experience higher absolute income losses, with median income losses of €2,000, compared to €1,000 for women. The corresponding means are €4,741.25 and €1,945.31 for self-employed men and women, respectively.

To measure relative losses, we relate the magnitude of income losses to 2019 earnings by dividing the absolute monthly losses in gross earnings by the monthly gross earnings of the previous year. However, since intra-year changes in income are fre-

[5]Note that the question on the precise amount of income losses was not, unfortunately, included in the first of the nine waves of the SOEP-COV. Since this happens to be the wave with the largest number of interviewees (see Figure 4.B.1), we are left with 104 self-employed individuals who reported income losses. We report both median and mean losses, but consider the median a superior statistic of centrality in this context, given that the distribution of income losses is, as expected, strongly right skewed (Sorgner et al., 2017).

Figure 4.3: The distributions of monthly relative losses in income (gross earnings) among self-employed individuals



(a) All (b) Gender differences

Note: Figures 4.3a and 4.3b display boxplots for relative monthly income losses among all self-employed individuals as well as self-employed men and women. The red marker indicates the median. The upper and lower end of the box display the range between the 25th and 75th percentiles. The whiskers span all data points within 1.5 inter-quartile range of the nearer quartile. Blue dots indicate observations outside the whiskers.

quent among the self-employed, the following results should be interpreted with some caution.[6]

The results for relative income losses are shown in Figure 4.3. Figure 4.3a displays the boxplot for all self-employed individuals. The median and mean of relative income losses among all self-employed individuals are 0.77 and 1.54, respectively. Figure 4.3b displays the boxplot for self-employed men and women. The median is 0.79 for self-employed women and 0.69 for men. Thus, in contrast to absolute losses, this suggests that the relative income losses tend to be larger for women. However, a formal median comparison indicates that we cannot reject equality of medians for self-employed men and women.

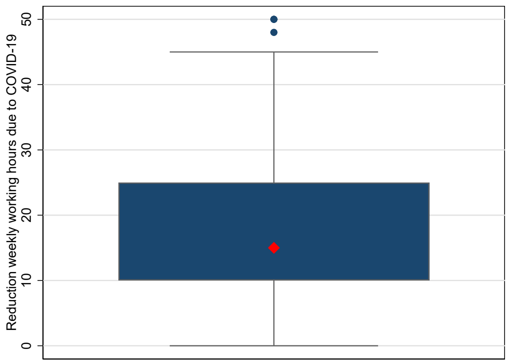
Turning to the reduction of weekly working hours due to COVID-19, we find that the median and mean absolute decreases are 15 and 18.07 hours, respectively.[7] The corresponding distribution is displayed in Figure 4.4a. Figure 4.4b shows that the median and mean reduction of working hours for self-employed men are 19 and 18.60 hours, respectively. The corresponding figures for self-employed women are slightly

[6] Compared to the previous analysis, we lose an additional 23 observations. The final sample includes 81 observations. Moreover, we do not adjust for inflation. However, first projections indicate that the inflation rate is approximately close to zero for 2020. Lastly, we partly observe individuals in different months, i.e. we are not able to account for seasonality. This possibly introduces some additional measurement error.

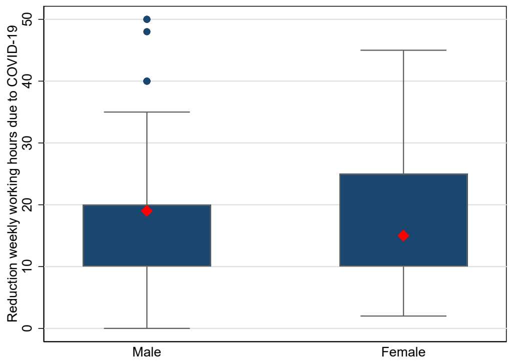
[7] We have information on reductions in working hours for all waves of the SOEP-CoV.

smaller, with a median of 15 and a mean of 17.61. Yet again, formal tests of equality across groups do not allow us to reject the hypothesis of no differences.

Figure 4.4: The distributions of the reduction of weekly working hours among the self-employed



(a) All



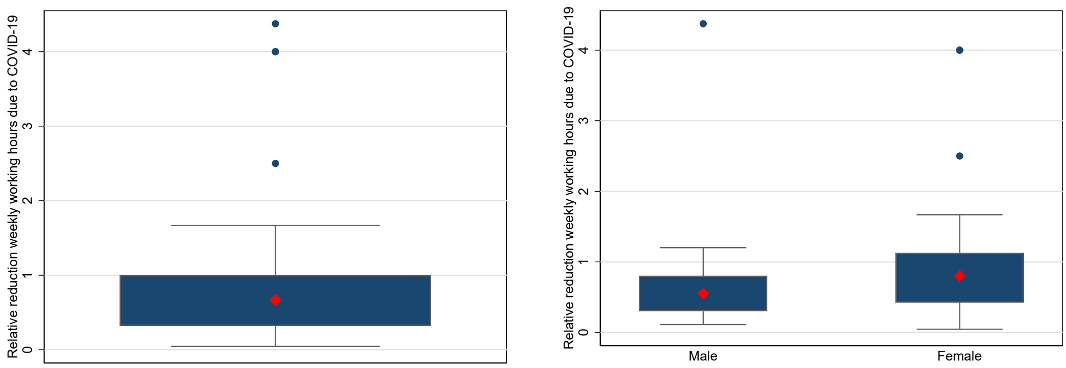
(b) Gender differences

Note: Figures 4.4a and 4.4b display boxplots for reductions in weekly working hours among all self-employed individuals as well as self-employed men and women. The red marker indicates the median. The upper and lower end of the box display the range between the 25th and 75th percentiles. The whiskers span all data points within 1.5 inter-quartile range of the nearer quartile. Blue dots indicate observations outside the whiskers.

Lastly, we focus on relative reductions in weekly working hours. We divide the decrease in weekly working hours due to COVID-19 by the actual weekly working hours of the previous year. The distributions are depicted in Figure 4.5.[8] Figure 4.5a displays the respective distribution for all self-employed individuals. The median and mean are 0.6 and 0.78, respectively. Figure 4.5b displays the corresponding gender-specific distributions. For self-employed men, the median and mean of relative working hours reductions are 0.5 and 0.77. For self-employed women, these figures are 0.63 and 0.79, respectively. Once again, the differences between men and women are not statistically significant.

[8]For the figures, we dropped a single observation with a relative reduction of 10.

Figure 4.5: The distributions of relative reductions of weekly working hours among the self-employed



(a) All (b) Gender differences

Note: Figures 4.5a and 4.5b display boxplots for relative reductions in weekly working hours among all self-employed individuals as well as self-employed men and women. The red marker indicates the median. The upper and lower end of the box display the range between the 25th and 75th percentiles. The whiskers span all data points within 1.5 inter-quartile range of the nearer quartile. Blue dots indicate observations outside the whiskers.

4.4 Multivariate analysis

Our descriptive results in the previous section show that the crisis following the COVID-19 pandemic impacts the female self-employed considerably more than all other groups. In this section, we perform multivariate analyses to better understand how these differences emerge.

4.4.1 Comparison of the self-employed and employees

To put the analysis of the gender gap among the self-employed into the larger context, we start with a comparison of all self-employed individuals with employees. Table 4.1 shows the results of a regression of indicators for a decrease in income, a decrease in working hours, and working from home, respectively, on an indicator for self-employment. While the odd-numbered columns only include state indicators as well as week indicators, the even columns expand the set of controls to include our complete set of controls.[9] With only state and week fixed effects as controls, self-employed individuals are 42 percentage points more likely to have experienced an income loss and 30 percentage points more likely to have experienced a reduction in working hours compared to employees. Self-employed individuals are also about six percentage points more likely to work from home.

The comparison of odd-numbered with even-numbered columns of Table 4.1 reveals that individual-level and household-level characteristics explain very little of the differences between self-employed individuals and employees with respect to the probability of income losses and hours reductions. The coefficient on the indicator for self-employment remains almost unchanged when adding controls (compare column (1) to column (2) and column (3) to column (4), respectively). Having a migration background appears to significantly increase the probability of suffering income losses and hours reductions, while a higher household income has the opposite effect. That is consistent with the finding of Fairlie (2020), who also finds a racial gap in how the self-employed are hit by the COVID-19 pandemic. By contrast, the probability of working

[9]Note that the estimates displayed in odd columns somewhat differ from the raw self-employment gap due to the inclusion of state and time fixed effects. The inclusion of these fixed effects is important since individuals in the sample were interviewed at different stages of the pandemic. Over time, the incidence of, for example, working from home changed. Thus, an accurate reflection of a self-employment (gender) gap requires that the dynamic of the pandemic is accounted for.

from home seems to be explained by the added controls: Individuals from more affluent households are more likely to be working from home during the pandemic, likely a result of selection into jobs that are more easily done from home (e.g. office jobs, see Alipour et al., 2020). Similarly, better-educated individuals are significantly more likely to work from home, so are parents. To pin down the relevance of industry fixed effects, Table 4.A.5 in the appendix displays the R-squared alongside the coefficients on the self-employment indicator for the unrestricted models in Table 4.1, both with and without the inclusion of industry fixed effects. The R-squared increases substantially once industry effects are accounted for, implying that industry-variation contributes significantly to explaining the respective outcomes.[10] However, differential selection into industries adds rather little to describing the overall differences between employees and the self-employed, as evidenced by the marginal changes in the self-employment gap once industry fixed effects are accounted for.

Since our observations do not seem to be driven by differences in characteristics, we then investigate whether differential associations of these characteristics with the outcome variables can explain the differential impact of the pandemic on the self-employed and employees. Therefore, we estimate our full model for each of our outcomes separately for both the self-employed and for employees. We also present p-values of Chow-tests comparing the coefficients across models.[11] Tables 4.A.6 to 4.A.8 in the appendix show the corresponding results.

With respect to the probability of an income decrease, it appears that the associations between individual-level characteristics and the outcomes differ only a little between the models for the self-employed and employees. There appears to be a differential relationship with respect to unemployment experience, which, however, seems to be relevant only for the self-employed (Table 4.A.6). With respect to the probability of a decrease in working hours, we again observe few differences between the models. Most notably, the presence of children in the household (school age or younger) increases the probability of a reduction in working hours by 21 percentage points for self-employed individuals while household size itself decreases the probability of a reduction in working hours by nearly eight percentage points on average. The latter

[10]Once industry effects are added, the R-squared increases by around 50% for the probability of working from home while nearly doubling for the probability of facing a decrease in income and hours, respectively.

[11]The p-values stem from a Chow-test after seemingly unrelated regressions.

might point to the presence of another helping individual in the household so that the self-employed individual is able to keep working. We do not observe comparable effects for employees (Table 4.A.7).[12]

Turning to the probability of working from home, we observe that older self-employed individuals are less likely to work from home, while there is no age gradient for employees (Table 4.A.8). Moreover, the correlation with household income as well as household size operates in opposite directions for self-employed and employed individuals. We also find some differences when it comes to personality traits, but also similarities: High scores in openness for experience increase the probability among both the self-employed and the employed to work from home. Conversely, it turns out that the observed strong and positive association between the probability of working from home and socio-economic status (income and education) is only true for employees, but not for the self-employed.

We then investigate the differences in the estimates of the industry fixed effects. Figure 4.B.2, shows the estimated fixed effects in increasing order of magnitude along with the associated 95% confidence intervals, separately for the self-employed and employees. The agricultural sector serves as the reference category (according to the Nomenclature of Economic Activities, NACE Rev. 2). For all outcomes, the point estimates are larger for the self-employed individuals. Moreover, the confidence intervals suggest a steeper gradient in the estimates of the fixed effects for the self-employed than for the employees throughout. Thus, it appears that differences in the variation of industry fixed effects between the self-employed and employees do contribute to the observable differences in the respective outcomes.

[12] Note that during the observation period, child-care facilities and schools were closed or only provided services for essential workers. Thus, a potential explanation for the differences could be that employees face stronger restrictions should they desire to reduce their working hours.

Table 4.1: Restricted and unrestricted model for difference of likelihood that income or working hours decreased or individual works from home between employees and self-employed respondents

	(1) Income	(2) Income	(3) Working hours	(4) Working hours	(5) Remote work	(6) Remote work
Self-employed	0.418***	0.421***	0.301***	0.302***	0.061**	0.021
	(0.029)	(0.031)	(0.029)	(0.031)	(0.030)	(0.032)
Demographics:						
Gender: Female		0.019		0.022		-0.013
		(0.013)		(0.016)		(0.017)
Age		0.006		-0.003		-0.005
		(0.005)		(0.005)		(0.005)
Age squared		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
Migration background		0.040**		0.040**		-0.026
		(0.016)		(0.019)		(0.019)
Big 5:						
Extraversion (2019)		0.000		0.008		-0.001
		(0.006)		(0.007)		(0.008)
Conscientiousness (2019)		-0.010		-0.018**		0.001
		(0.007)		(0.008)		(0.008)
Openness to experience (2019)		0.010		0.006		0.025***
		(0.006)		(0.007)		(0.008)
Neuroticism (2019)		-0.004		0.001		-0.008
		(0.006)		(0.007)		(0.007)
Agreeableness (2019)		0.004		-0.004		0.002
		(0.006)		(0.007)		(0.008)
Household context:						
HH Size (2019)		0.006		0.011		-0.008
		(0.007)		(0.008)		(0.009)
Married		0.021		0.016		-0.021
		(0.015)		(0.017)		(0.018)
School child or younger		0.007		-0.004		0.049**
		(0.018)		(0.021)		(0.022)
Log. of HH net income (2019/18)		-0.039**		-0.034*		0.098***
		(0.016)		(0.018)		(0.020)
Education (ref. low):						
Intermediate education		0.031		0.023		0.073***
		(0.019)		(0.022)		(0.020)
High education		0.011		-0.005		0.293***
		(0.021)		(0.024)		(0.024)
Unemployment experience		0.000		0.005*		-0.005**
		(0.003)		(0.003)		(0.002)
Mean of outcome	0.169	0.169	0.222	0.222	0.395	0.395
Observations	3,531	3,531	3,518	3,518	3,533	3,533
R^2	0.11	0.23	0.05	0.13	0.03	0.31

Note: Table 4.1 displays models with and without controls for differences between self-employed and employees. All models include state and week fixed effects. Columns (1), (3) and (5) display results for the models without controls. Columns (2), (4) and (6) display results for the models with controls. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

In summary, it seems that the differential impact of the COVID-19 pandemic between employees and the self-employed with respect to income and working hours is neither primarily driven by differences in individual- and household-level characteristics nor by selection into different industries, but by differences in the association of these factors with the respective outcomes. The pandemic shock hit the self-employed uniformly harder. This seems plausible as employees are often shielded from job and income losses by employment contracts and job protection legislation, while such mechanisms do not exist for the self-employed. By contrast, individual- and household-level characteristics can nearly fully account for differences in the likelihood of working from home between self-employed and employed individuals.

Thus far, we focus our analysis on the population of (self-)employed individuals in 2020. However, employees may have lost their job over the course of the pandemic and self-employed individuals may have terminated their business. To account for this, we look at the working population of 2019 and investigate whether individuals who were self-employed in 2019 differ from those who were employees with respect to the probability of changes in income, changes in working hours, and job loss. The latter is defined as the proportion of individuals who transitioned into non-employment between 2019 and 2020 and who respond that this transition was due to the COVID-19 pandemic. The results are shown in Table 4.A.9 in the appendix. Overall, 1.7% of those working in 2019 are non-employed in 2020 because of the pandemic. Importantly, self-employed individuals are 1.2 percentage points more likely to have terminated their business than employees are to have lost their job, although this difference is not statistically significant. Note as well that the reported results for income and working hours changes slightly differ from those in Table 4.1. This is explained by the focus on the employment status of 2019, rather than 2020 in Table 4.A.9. Differences result from two sources: First, employees surveyed in 2019 may have become self-employed between the times of the interview in 2019 and 2020, and vice versa. Second, individuals who were not in employment at the time of the interview in 2019 may have founded a business prior to the time of the interview in 2020. However, the differences in the reported results between Table 4.1 and Table 4.A.9 are minor.

4.4.2 Gender differences among the self-employed

As discussed in Section 4.3.3, we observe considerable gender differences in the probability of income declines among the self-employed. Section 4.4.1 further reveals that self-employed individuals are, in general, much more likely to suffer income losses than employees. Turning to our core analysis, we investigate how self-employed as well as employed women are affected by the COVID-19 pandemic in comparison to their male counterparts. We apply the Gelbach (2016) decomposition to further analyze the gender differences with respect to the likelihood of a decline in income due to the COVID-19 pandemic. This decomposition reveals the individual contributions of covariates to the gender gap, thus assigning each covariate-bundle a proportion of the overall contribution. Importantly, it is not path dependent, as this decomposition is, unlike sequential covariate addition, invariant to the sequence in which we would usually insert the covariates to gauge the stability of the coefficient of interest. In our analysis, the Gelbach decomposition answers the question of how much of the change in the gender gap can be attributed to different variables in the set of controls as we move from the base specification, the restricted model, to the full specification that includes all controls, the unrestricted model (for more details on the methodology see Appendix 4.C).

In our sample of self-employed individuals, we observe a gender gap of 17.4 percentage points in the likelihood of experiencing an income loss in our restricted model. This can be inferred from column (1) in Table 4.2.[13] Relative to self-employed men, self-employed women are 36.9% more likely to experience an income loss because of the COVID-19 pandemic. As discussed in Section 4.3.3 and confirmed in Table 4.A.10 in the appendix, there is no comparable gender gap among employees. In our unrestricted model in column (2) of Table 4.2, the gender gap decreases to 8.1 percentage points and is statistically indistinguishable from zero. This outcome implies that our controls can explain about 9.3 percentage points, or 53.4%, of the initial gender gap.[14]

The largest share of the gender gap in income losses can be explained by the fact

[13]Once again, the estimates displayed in the restricted models somewhat differ from the raw gender gap due to the inclusion of state and time fixed effects.

[14]The corresponding analysis of the magnitude of earnings losses are relegated to Section 4.D in the appendix. Since sample sizes decrease considerably, the analysis suffers from imprecision. Effect sizes still confirm our main conclusions, even for the changes at the intensive margin.

that women are over-represented in industries in which individuals are more likely to experience income losses. This is seen in Figure 4.6a, which displays the results of the Gelbach decomposition: 9.2 percentage points, or 98.8% of the total change, can be explained by NACE fixed effects.[15] Demographic characteristics, particularly age, explain as much as 33.8% of the total change in the gender gap between the unrestricted and restricted models. Other groups of characteristics add nearly nothing to the total change in the gender gap.[16]

Thus, the industry-specific likelihood of an income loss is positively associated with the share of women in the respective industry. In Figure 4.7, we display binned scatter plots for the association between the respective industry-specific fixed effects in the likelihood of an income loss and share of women for self-employed individuals and employees, respectively.[17] We observe a positive association between the industry fixed effects and the share of women in the respective industries. The OLS coefficient for the underlying relationship implies that a 10 percentage point higher share of women in a given industry is associated with an increase in the likelihood of experiencing an income loss of about 5.6 percentage points.

Moreover, the results in columns (3) and (5) of Table 4.2 do not support the notion of a gender gap in the likelihood of a decline in working hours and working from home.[18][19] However, the change in the OLS coefficient for the indicator for being female between the restricted and unrestricted model and Figure 4.6b suggests an economically significant change in the likelihood of a decline in working hours of about 11.9 percentage points, which is more

[15]Detailed results of the Gelbach decomposition are depicted in Table 4.A.11.

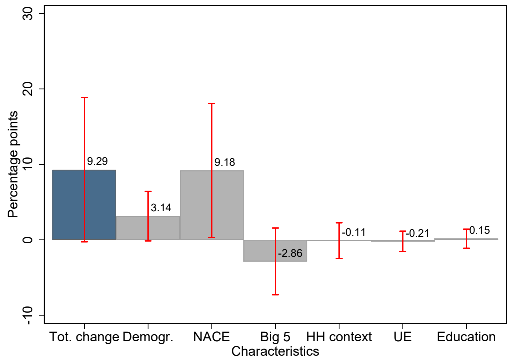
[16]Figure 4.B.3 shows the decomposition for employees corresponding to Table 4.A.10.

[17]In Figure 4.7, we calculate the share of women in the respective industries over the complete working sample, i.e. we do not distinguish between self-employed and employed individuals.

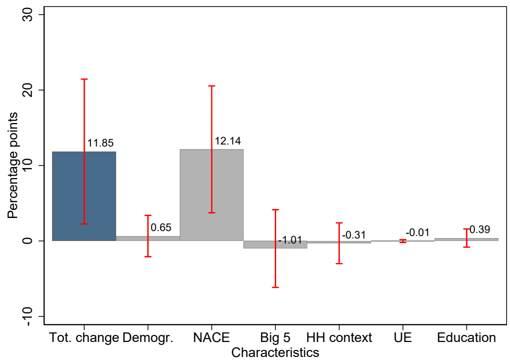
[18]Note that the possibility to work from home has a moderating effect on the likelihood of an income and working hours decline. If we include the indicator for working from home in the unrestricted models, the coefficients suggest that working from home is associated with a 13.72 ($p = 0.063$) and 15.62 ($p = 0.041$) percentage point lower likelihood of a decline in income or working hours, respectively. However, the Gelbach decomposition suggests that working from home because of the COVID-19 pandemic does not contribute to the gender difference in these two outcomes. These results are available upon request.

[19]There might also exist initial gender differences in the standard workload among the self-employed. Therefore, we also examined whether including actual weekly working hours of the previous year alters the estimated gender gaps in a meaningful way, which is not the case.

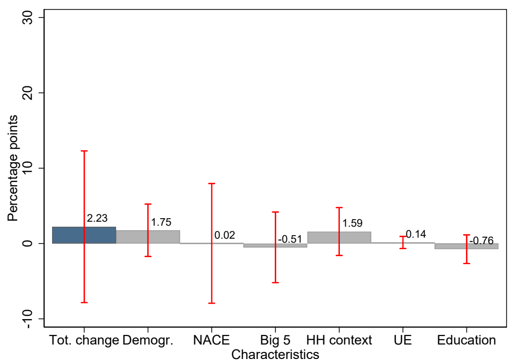
Figure 4.6: Gelbach decomposition of the gender gap in labor market outcomes among self-employed respondents



(a) Likelihood of income decline



(b) Likelihood of decline in working time



(c) Likelihood of remote work

Note: Figures 4.6a to 4.6c display the Gelbach decomposition of the gender gap in the likelihood of an income and working time decline among self-employed respondents. The reference category for the industry indicators is “Crop and animal production, hunting and related service activities.” Red bars indicate 95% confidence intervals based on robust standard errors.

Table 4.2: Restricted and unrestricted model for likelihood that income or working hours decreased or individuals are working from home among self-employed individuals

	(1) Income	(2) Income	(3) Working hours	(4) Working hours	(5) Remote work	(6) Remote work
Gender: Female	0.174***	0.081	0.068	-0.051	-0.017	-0.040
	(0.058)	(0.073)	(0.060)	(0.073)	(0.057)	(0.069)
Demographics:						
Age		0.027		0.007		-0.042**
		(0.019)		(0.020)		(0.021)
Age squared		-0.000*		0.000		0.000*
		(0.000)		(0.000)		(0.000)
Migration background		0.064		0.120		-0.117
		(0.110)		(0.099)		(0.085)
Big 5:						
Extraversion (2019)		0.011		0.067*		0.046
		(0.040)		(0.037)		(0.037)
Conscientiousness (2019)		-0.031		-0.058		0.033
		(0.039)		(0.038)		(0.037)
Openness to experience (2019)		0.066*		0.051		0.058*
		(0.038)		(0.036)		(0.034)
Neuroticism (2019)		-0.031		-0.003		-0.013
		(0.036)		(0.039)		(0.035)
Agreeableness (2019)		-0.040		-0.067*		-0.032
		(0.035)		(0.034)		(0.033)
Household context:						
HH Size (2019)		-0.061		-0.076**		0.092***
		(0.039)		(0.036)		(0.033)
Married		0.037		-0.010		0.026
		(0.073)		(0.078)		(0.071)
School child or younger		0.045		0.211**		-0.018
		(0.103)		(0.094)		(0.101)
Log. of HH net income (2019/18)		-0.026		0.100*		-0.146***
		(0.058)		(0.058)		(0.052)
Education (ref. low):						
Intmermediate education		-0.102		0.074		-0.108
		(0.125)		(0.114)		(0.112)
High education		-0.149		-0.026		0.057
		(0.132)		(0.120)		(0.119)
Unemployment experience		-0.026**		0.001		-0.013
		(0.012)		(0.010)		(0.011)
Mean of outcome	0.552	0.552	0.495	0.495	0.457	0.457
Observations	310	310	309	309	311	311
R^2	0.13	0.41	0.09	0.40	0.16	0.47

Note: Table 4.2 displays restricted and unrestricted models underlying the Gelbach decomposition. All models include state and week fixed effects. Columns (1), (3) and (5) display results for the restricted models. Columns (2), (4) and (6) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

than fully accounted for by the fact that, again, women are disproportionately represented in those industries hardest hit by the COVID-19 pandemic. In addition, Figure 4.7c suggests a positive association between the share of women across industries and the likelihood of experiencing a decline in working hours in these industries. This constitutes evidence that the industry affiliation moderates the relationship between the likelihood of a decline in working hours and the gender of self-employed respondents, while there is no evidence for such a relationship on the probability of working from home. We also do not find support for such a relationship among employees. Table 4.A.10 and 4.A.12 together with Figure 4.B.3 in the appendix and the binned scatter plots for employees in Figure 4.7 support this conclusion.

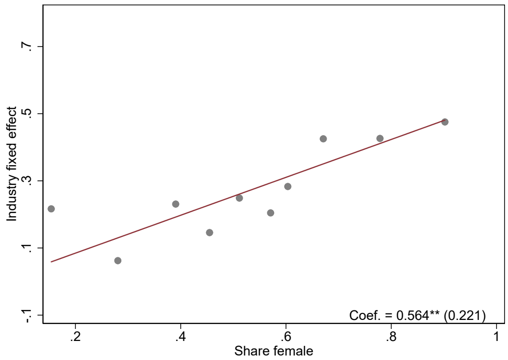
In Table 4.A.13 in the appendix, we display the five industries with the highest and lowest shares of women, respectively. The industries with the highest share of women include, for example, the hospitality sector and personal services – industries that were hit particularly hard by the COVID-19 crisis.[20] For each of these industries we also show the associated industry fixed effect corresponding to column (2) of Table 4.2. The average share of women for these industries in our sample is 82.57% and the average estimate of the fixed effects is 0.41.[21] Conversely, the average share of women in the five industries with the lowest shares of women in our sample is 25.68% and the average fixed effect for these industries is 0.19.[22] Thus, the contribution of industry fixed effects to the likelihood of suffering income losses due to the COVID-19 pandemic is largest in industries where women are over-represented.

[20]In our sample, the industries with the highest shares of women are, in decreasing order, "Other personal service activities," "Social work activities without accommodation," "Retail trade, except of motor vehicles and motorcycles," "Accommodation," and "Human health activities." A detailed breakdown of industries is limited by sample size restrictions. In Table 4.A.13 we only display industries with at least ten observations.

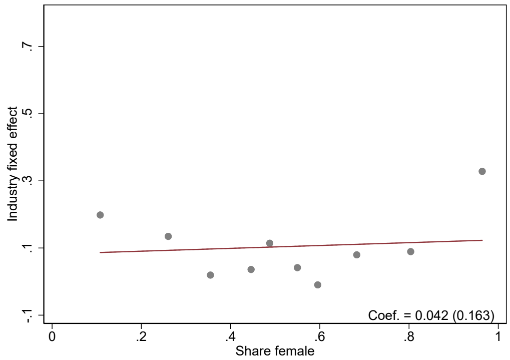
[21]The reference category is the agricultural sector.

[22]The five industries with the lowest shares of females are, in increasing order, "Land transport and transport via pipelines," "Printing and reproduction of recorded media," "Specialized construction activities," "Computer programming, consultancy and related activities," and Manufacture of machinery and equipment n.e.c."

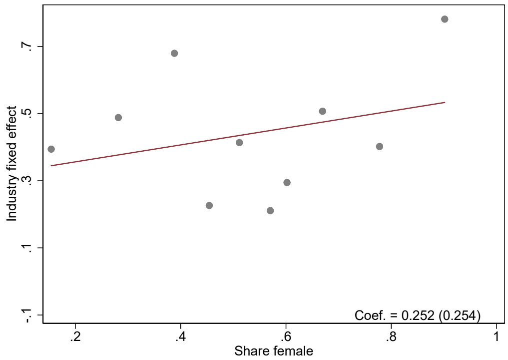
Figure 4.7: The association between industry specific fixed effects for the probability of an income or working time decrease as well as the probability of working from home and the share of women in the respective industry



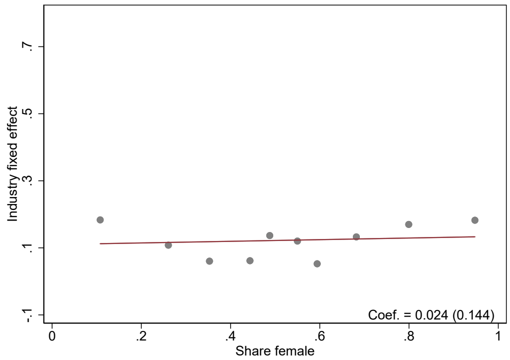
(a) Income decline for self-employed



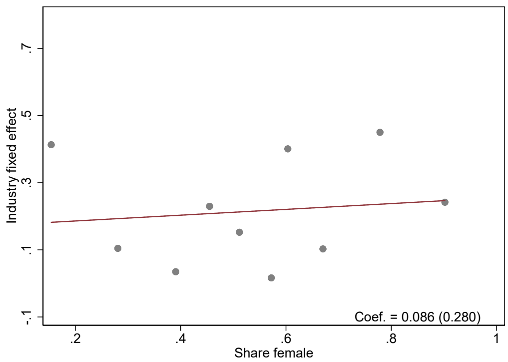
(b) Income decline for employees



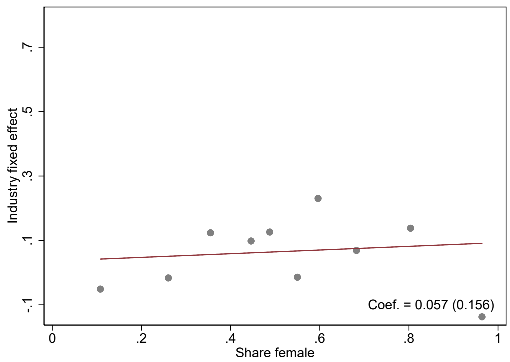
(c) Working time decline for self-employed



(d) Working time decline for employees



(e) Remote work for self-employed



(f) Remote work for employees

Note: Figures 4.7a to 4.7f display the association between industry specific fixed effects and the share of women in the respective industry for the working population in 2020. The fixed effects stem from a regression of our three outcomes on industry indicators, respectively. The share of women corresponds to the share of women in the respective industry in our working sample. Both figures correspond to a binned scatterplot. The regression coefficients stem from an OLS regression of the industry fixed effects on the share of women in the respective industries. Robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

4.5 Potential mechanisms

In this section, we investigate potential mechanisms driving our results. Note that the gendered industry effects presented in Section 4.4.2 encompass a variety of factors: Not only do they suggest the existence of direct effects of the pandemic that impact industries to varying degrees, i.e. through government-imposed restrictions, but also the importance of other NPIs, such as the closure of schools or day-care centers, and the indirect effects these have on income or hours worked, i.e. through changes in the intra-household allocation of time. Therefore, the overall contribution of the industry fixed effects is the product of the strength of the selection into industries as well as the association of the respective industry with the respective outcome.

In the following, we further characterize these relationships. We investigate to what extent direct regulations, or shortages in supply or demand, drive the disproportionate impact of the COVID-19 pandemic on self-employed women. We then test whether gendered specialization in home production might have contributed to the differential impact of the COVID-19 pandemic among self-employed women and men.

4.5.1 Business-related distortions due to the COVID-19 pandemic

In the SOEP-CoV questionnaire, self-employed respondents were asked whether they have been affected by several events in the wake of the COVID-19 pandemic and associated NPIs. Of these, we focus on events that might have detrimental effects on the self-employed respondents' income or working time. These are "being affected by regulations, e.g. opening hours" (Restrictions), "suppliers are not able to deliver parts or products to perform business" (Supply), and "customers are cancelling services or orders" (Demand). We apply the Gelbach decomposition to decompose the gender gap in the likelihood that the self-employed respondents report to have been affected by these events. Table 4.3 displays the restricted and unrestricted model for these three events.[23]

We find that self-employed women are 20.2 percentage points more likely than their

[23] See Table 4.A.2 and Table 4.A.3 for summary statistics on the dependent variables used in this section.

male counterparts to state that they are affected by rules or restrictions. We do not find such differences for the supply of intermediate goods or for demand shortages. In Figure 4.8, we show detailed Gelbach decompositions of the gender gap for business-related events. The Gelbach decomposition in Figure 4.8a, along with the results in Table 4.3, provide evidence that it is, once again, the disproportionate representation of women in industries most affected by the pandemic that explains the differential effects.[24] Our full set of covariates explains about 15 percentage points of this gender gap, with about 9 percentage points thereof attributable to industry fixed effects. While the total change of the gender gap between the restricted and unrestricted model is significant at the five percent level of significance, the contribution of industry fixed effects is significant at the ten percent level of significance.

Moreover, we find that government-imposed restrictions contribute significantly to the gender gap in the likelihood of an income decline. This is shown in Figure 4.9, where we include indicators for the three business-related events in the wake of the COVID-19 pandemic in the Gelbach decomposition of the gender gap in income losses.[25] Among the three business-related events considered, being affected by rules and restrictions due to the COVID-19 pandemic is the only relevant contributor to the gender gap in income loss. As depicted in Figure 4.9a, rules and restrictions account for 4.5 percentage points of the total change of 10.3 percentage points.[26] At the same time, the contribution of industry fixed effects is considerably attenuated from 9.2 to 7.1 and is significant at the 10% level of significance, suggesting that government-imposed restrictions disproportionately affect industries in which women are over-represented and that those restrictions contribute to positively to the likelihood of an income decline.

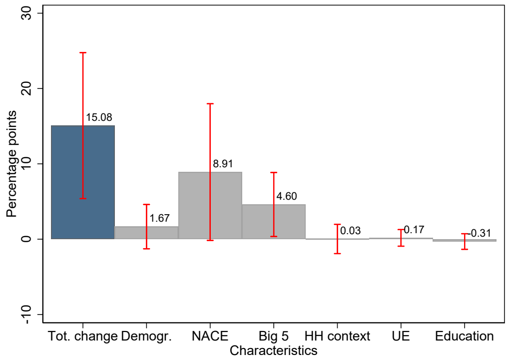
[24]The detailed results for the Gelbach decomposition are depicted in Table 4.A.14 in the appendix.
[25]For the sake of brevity, we consolidate all other characteristics in the category "Remainder."
[26]Detailed results for the Gelbach decomposition are displayed in Table 4.A.15 in the appendix.

Table 4.3: Restricted and unrestricted model for likelihood that business was affected by event

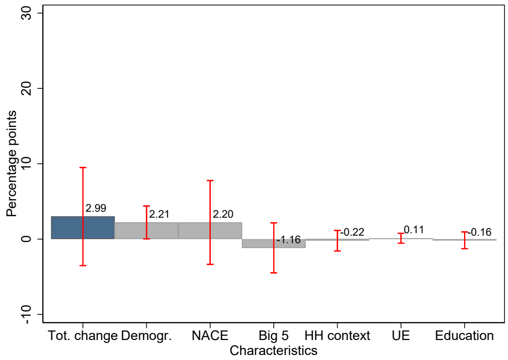
	(1) Restrictions	(2) Restrictions	(3) Supply	(4) Supply	(5) Demand	(6) Demand
Gender: Female	0.202***	0.051	-0.027	-0.057	0.052	-0.007
	(0.058)	(0.068)	(0.041)	(0.048)	(0.059)	(0.073)
Demographics:						
Age		-0.005		0.028**		0.022
		(0.019)		(0.013)		(0.019)
Age squared		0.000		-0.000**		-0.000*
		(0.000)		(0.000)		(0.000)
Migrant		0.092		0.014		0.032
		(0.090)		(0.075)		(0.097)
Big 5:						
Extraversion		0.039		-0.004		0.039
		(0.037)		(0.029)		(0.039)
Conscientiousness		-0.025		0.021		-0.046
		(0.036)		(0.024)		(0.039)
Openness		-0.030		-0.009		0.055
		(0.037)		(0.027)		(0.038)
Neuroticism		0.064*		-0.001		0.001
		(0.035)		(0.024)		(0.039)
Agreeableness		0.037		-0.038		-0.017
		(0.035)		(0.026)		(0.037)
Household context:						
HH Size		-0.001		0.024		-0.035
		(0.032)		(0.027)		(0.040)
Married		-0.019		-0.058		-0.041
		(0.073)		(0.056)		(0.079)
School child or younger		-0.091		-0.099		-0.038
		(0.096)		(0.078)		(0.108)
Log. HH net income		-0.057		0.015		0.018
		(0.057)		(0.044)		(0.060)
Education (ref. low):						
Intermediate education		-0.110		-0.147		-0.112
		(0.105)		(0.098)		(0.116)
High education		-0.054		-0.132		-0.100
		(0.108)		(0.103)		(0.120)
Unemployment experience		-0.016		-0.011**		-0.021**
		(0.011)		(0.005)		(0.009)
Mean of outcome	0.457	0.457	0.122	0.122	0.434	0.434
Observations	311	311	311	311	311	311
R^2	0.13	0.46	0.05	0.31	0.09	0.38

Note: Table 4.3 displays restricted and unrestricted models underlying the Gelbach decomposition for business events. All models include state and week fixed effects. Columns (1), (3) and (5) display results for the restricted models. Columns (2), (4) and (6) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

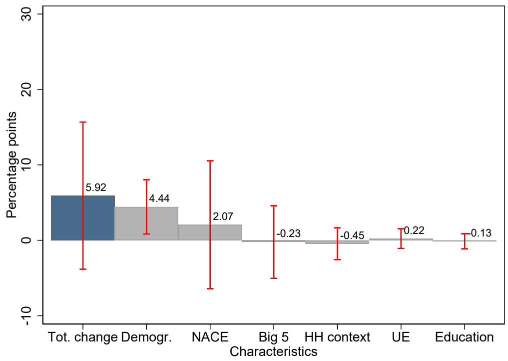
Figure 4.8: Gelbach decomposition of the gender gap in business-related events



(a) Rules or restrictions



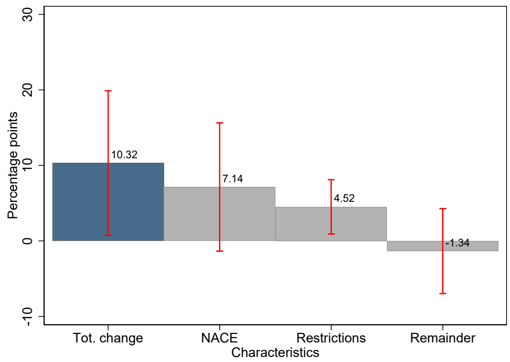
(b) Supply of intermediate products



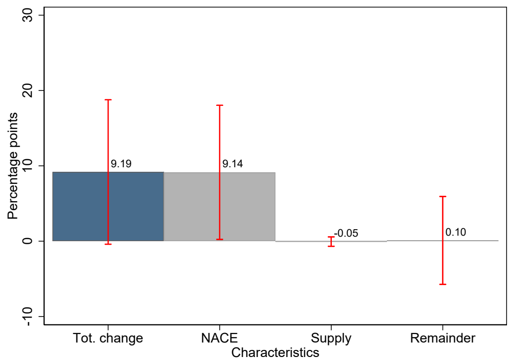
(c) Demand shortage

Note: Figures 4.8a to 4.8c display the contribution of the industry affiliation to the gender gap in various business-related events. Red bars indicate 95% confidence intervals and are based on robust standard errors.

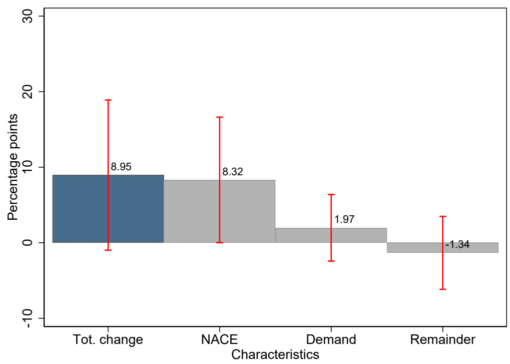
Figure 4.9: Contribution of the business-related events to the gender gap in the likelihood of an income decline



(a) Rules or restrictions



(b) Supply of intermediate products



(c) Demand shortage

Note: Figures 4.9a to 4.9c display the importance of various business-related events for the gender gap in the likelihood of an income decline. We summarized the residual characteristics in the category "Remainder." Red bars indicate 95% confidence intervals and are based on robust standard errors.

4.5.2 Household income and household chores

As noted previously, direct regulations of businesses are not the only government interventions that can potentially affect labor market outcomes of self-employed individuals. Other NPIs include the closure of schools and child-care facilities, which may also contribute to the observed gender gap. Assume that households maximize income subject to a time constraint. Further, assume decreasing returns and comparative advantages in household and market production, respectively. Under these conditions, both spouses would participate in the labor force in normal times. However, their respective contributions to the household income would be determined by their relative productivity in home and market production (e.g. Weiss, 1993; Bertrand et al., 2015). In this class of models, the partner who is relatively more productive at home production tends to spend more time with household chores or childcare. At the same time, their spouse spends more time in market production, where they are hypothesized to be relatively more productive, and thus earn a higher income.[27]

Given these assumptions, households need to re-optimize if, for instance, child-care facilities close. Then it is likely that the partner with the higher relative productivity in home production reduces time in market production while the other partner increases hours worked, *ceteris paribus*. One implication of this simplified model is that, if women tend to be the partner who is relatively more productive in home production, we would observe a gender gap in income and time decreases as a consequence of NPIs reducing the share of home production that can be outsourced, i.e. the closure of childcare facilities.[28] So far, we accounted for this by controlling for the presence of children and household size.

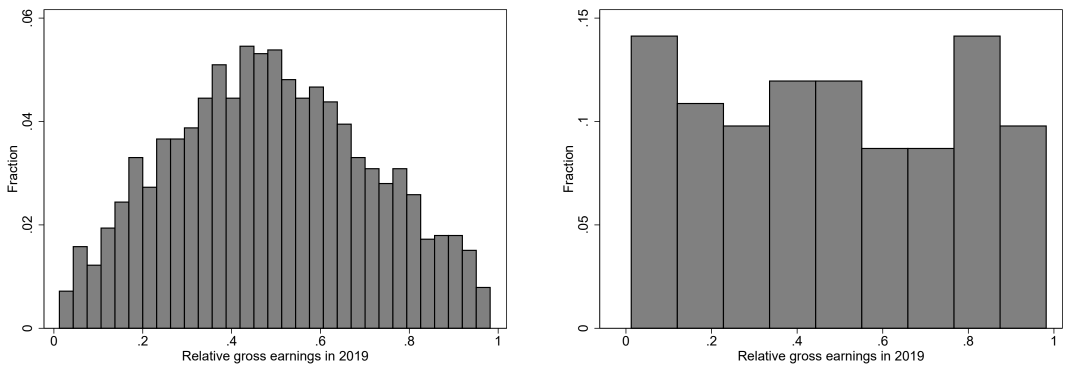
We now test this prediction by including an individual's earnings, relative to the overall earnings of the household, in our models. The concept is focal in the literature on gender norms (e.g. Bertrand et al., 2015; Foster and Stratton, 2021). For each respondent, we know the partner from 2018. Thus, we are able to link the partners' earnings from 2019 to each respondent. Then we calculate the relative earnings of each individual within each of these household pairs. Note that not every individual

[27] At this point, we abstract from gender norms, which could also explain the gendered response to a closure of child-care facilities.

[28] We are very grateful to an anonymous referee who suggested the discussion of intra-household dynamics.

in our data has a partner. In such cases, the relative earnings for this observation is 100% or 0%. We account for these single households via the inclusion of an indicator for having a partner in 2018. If an individual did not work in 2019, we impute zero earnings. The distribution of relative earnings is displayed in Figure 4.10. We then include relative earnings in the Gelbach decomposition. If the conjecture above is true, we would expect that women are more likely to have lower relative earnings and relative earnings would be negatively associated with the incidence of a decrease in working time, income, or the likelihood of working from home.

Figure 4.10: Distribution of relative earnings in 2019



(a) All (b) Self-employed

Note: Figures 4.10a and 4.10b display the relative earnings in 2019. In these figures, we discard observations for whom the relative earnings is zero or one.

With respect to the likelihood of income reductions we find some evidence for the first part of the conjecture. That is, the results indicate that households optimize and exploit comparative advantages. Table 4.4 displays the restricted and unrestricted model for our outcome variables. In addition to the standard set of controls, we now include the individual's share of household earnings in 2019. In addition, all models include an indicator for the presence of a partner. For the likelihood of an income decline due to the COVID-19 pandemic, the earnings share of the individual is significant at the ten percent level of significance. The point estimate suggests that a ten percent increase in the individual's earnings share is associated with a 2.6 percentage points reduction in the likelihood of an income reduction. Similarly, the Gelbach decomposition in Figure 4.11a suggests that women account for a smaller share of the total household earnings, on average, and that the share of household earnings is negatively associated with the likelihood of an income reduction due to the COVID-19 pandemic. This relationship accounts for 25.8% of the initial gender gap. However, the estimate is not very precisely estimated, meaning we cannot reject the hypothesis that this contribution is different from zero ($p = 0.104$). However, it is worth emphasizing that the gender gap almost completely vanishes as soon as we account for relative earnings (compare column (2) of Table 4.2 to column (2) of Table 4.4).

With respect to the likelihood of a reduction in working hours or the incidence of working from home, we find no evidence for a significant association with the individuals' earnings share within the household. The Gelbach decompositions in Figure 4.11b and Figure 4.11c likewise do not provide an indication that the relative income position contributes to explaining the gender gap. One interpretation of these findings is that a negative association would appear only for outcomes that translate directly into material well-being. For working time, this is not clear a priori. For self-employed individuals, there are various possible circumstances where working time reductions do not necessarily translate into reduced earnings. With respect to the incidence of working from home, other factors are likely more relevant, i.e. the extent to which the job of the self-employed individual or their partner can be performed remotely.[29]

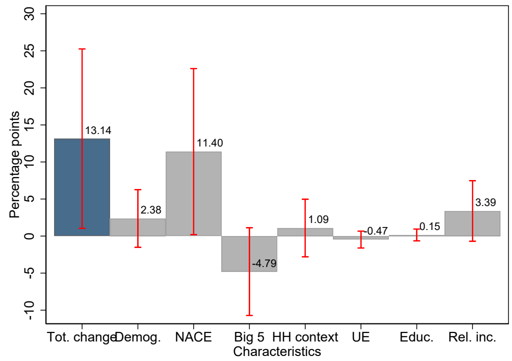
[29]We abstract from leisure in this analysis since we assume that individuals shift their time from market production to household chores.

Table 4.4: Restricted and unrestricted model for difference of likelihood that income or working hours decreased, accounting for relative income differences

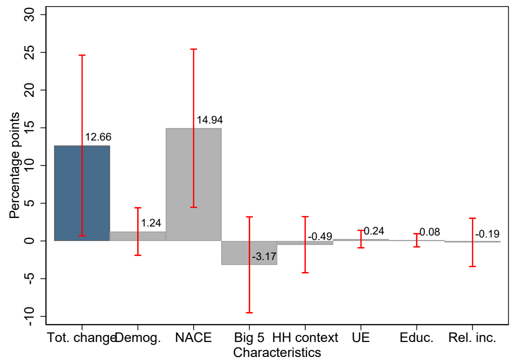
	(1) Income	(2) Income	(3) Working hours	(4) Working hours	(5) Remote work	(6) Remote work
Gender: Female	0.154**	0.022	0.066	-0.054	0.002	-0.029
	(0.067)	(0.086)	(0.070)	(0.088)	(0.066)	(0.077)
Demographics:						
Age		0.056*		0.016		-0.031
		(0.032)		(0.033)		(0.034)
Age squared		-0.001*		0.000		0.000
		(0.000)		(0.000)		(0.000)
Migration background		0.018		0.025		-0.207*
		(0.131)		(0.114)		(0.113)
Big 5:						
Extraversion (2019)		0.044		0.054		0.052
		(0.047)		(0.045)		(0.044)
Conscientiousness (2019)		-0.040		-0.016		-0.016
		(0.046)		(0.045)		(0.045)
Openness to experience (2019)		0.055		0.035		0.048
		(0.048)		(0.046)		(0.041)
Neuroticism (2019)		-0.062		-0.042		-0.015
		(0.042)		(0.044)		(0.040)
Agreeableness (2019)		-0.087**		-0.073*		-0.023
		(0.043)		(0.043)		(0.041)
Household context:						
HH Size (2019)		-0.072		-0.065		0.100***
		(0.050)		(0.043)		(0.036)
Married		0.072		-0.012		0.028
		(0.124)		(0.151)		(0.117)
School child or younger		0.056		0.247**		0.078
		(0.124)		(0.110)		(0.124)
Log. of HH net income (2019/18)		-0.064		-0.127**		-0.127**
		(0.069)		(0.066)		(0.064)
Education (ref. low):						
Intermediate education		0.019		0.090		-0.049
		(0.146)		(0.137)		(0.137)
High education		-0.033		0.064		0.065
		(0.161)		(0.142)		(0.149)
Unemployment experience		-0.025		0.013		-0.048***
		(0.020)		(0.021)		(0.017)
Income share		-0.260*		-0.002		0.143
		(0.135)		(0.156)		(0.136)
Mean of outcome	0.561	0.561	0.496	0.496	0.496	0.496
Observations	239	239	238	238	238	238
R^2	0.17	0.50	0.13	0.48	0.13	0.48

Note: Table 4.4 displays restricted and unrestricted models underlying the Gelbach decomposition. All models include state and week fixed effects as well as indicators for having a partner. Columns (1), (3) and (5) display results for the restricted models. Columns (2), (4) and (6) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

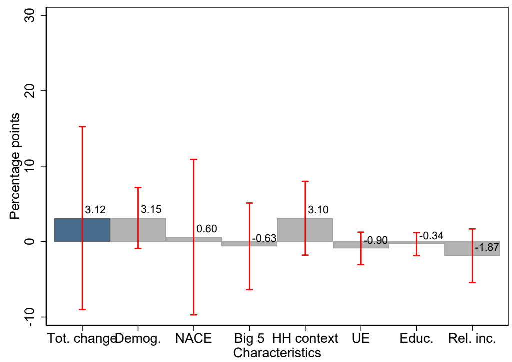
Figure 4.11: Gelbach decomposition of the likelihood of an income or working time reduction, testing the specialization in the household context



(a) Reduction of income



(b) Reduction in weekly working hours



(c) Remote work

Note: Figures 4.11a to 4.11c display the Gelbach decomposition of the gender gap in the likelihood of income reduction, reduction in working time, and working from home among self-employed respondents. Red bars indicate 95% confidence intervals based on robust standard errors.

4.6 Conclusion

We analyze how the economic shock related to SARS-CoV-2 affects the self-employed in comparison to employees, and focus in particular on the female working population. We show that the 4.2 million self-employed men and women are 42 percentage points more likely to experience an income loss than employees and that they have a 30 percentage points higher chance of experiencing a decrease in working hours. This differential impact on the two employment forms cannot be explained by differences in individual-level characteristics or selection into different industries. The self-employed are more likely to suffer income losses and reductions in working hours throughout.

Unlike for self-employed workers, employees' wages and working hours in Germany are more rigid than in comparable countries. In addition, to prevent mass layoffs, the German government expanded "*Kurzarbeit*," its well-established short-time work scheme that allows for temporary reductions in the wages and hours of employees. Indeed, the fraction of employees who experience income losses is proportional to the fraction of employees in short-time work schemes (Kritikos et al., 2020). Thus, it appears that the labor market impact of the COVID-19 pandemic was mitigated by *Kurzarbeit*.

Among the self-employed, we find that women are about one-third more likely to face income losses due to the COVID-19 pandemic than men. We do not find a comparable gender gap among employees, which is likely a result of labor market rigidities. Our results reveal that the largest share of gender differences among the self-employed is attributable to the fact that self-employed women work disproportionately in industries that are more severely affected by the COVID-19 pandemic than men. This is supported by the observable gender gap in the extent to which self-employed individuals were affected by government-imposed restrictions, such as the regulation of opening hours. We provide evidence that this directly translates into gender differences in income losses. Moreover, we find suggestive evidence that gendered household production also contributes to the gender gap in income losses. Still, this is of second order compared to the contribution of industry effects.

Our study has important policy implications that may be applicable for policy responses to the further development of the current pandemic or for future pandemics

(Petrovan et al., 2020). We show that the self-employed, in particular women, are hit significantly harder by this systemic shock than other parts of the working population, which is, in part, a direct consequence of policy measures enacted to contain the spread of the virus. This outcome should also be seen in the context of the slowly increasing willingness of women to enter self-employment. If self-employed women feel less supported by policy measures during such a systemic shock than female employees, society risks that they will start turning away from this employment form. Thus, the gender gap in self-employment may widen again. This could negatively affect growth, notably in parts of the economy that depend strongly on female self-employment. The design of policy measures intending to mitigate negative economic shocks in the ongoing or in comparable future crisis situations, should, therefore, account for this variation in economic hardship. Given our finding that government-imposed restrictions are a factor through which this unequal impact of the pandemic emerges, targeted policies that restore gender equity seem particularly relevant.

Given our finding that the self-employed are disproportionately affected by the COVID-19 pandemic, policy makers may consider different measures aimed at supporting them. Every such policy measure involves the risk of moral hazard. That is, it provides incentives for self-employed individuals to engage in risky behavior in a way it would not occur in the absence of support schemes. On the other hand, the detrimental effect of the COVID-19 pandemic on the self-employed is not the result of individual decision-making, rather it is a systematic and unexpected shock, and in part a direct consequence of government regulation. More generally, any support scheme for the self-employed may create both negative and positive externalities, which are to be weighted against each other. For instance, self-employment and entrepreneurship are shown to have a positive effect on growth (Stoica et al., 2020). As such, support schemes which successfully retain the propensity to remain self-employed through the crisis have the potential to facilitate recovery after the COVID-19 pandemic.

Appendix

4.A Additional tables

Table 4.A.1: Variable descriptions

(1) Variable	(2) Description	(3) Year of origin
Income (gross) decrease	Indicator reflecting decrease of monthly gross income decrease due to COVID-19 pandemic.	2020
Working hour decrease	Indicator reflecting decrease of weekly working hours decrease due to COVID-19 pandemic.	2020
Income loss	Exact amount of lost income due to COVID-19 pandemic.	2020
Number of working hour decrease	Exact number for the decrease of weekly working hours due to COVID-19 pandemic.	2020
Remote work	Indicator reflecting working from home due to COVID-19 pandemic.	2020
Age	Difference between survey year and birth year.	pre 2020
Female	Indicator for being female.	pre 2020
Migration background	Indicator for having direct or indirect migration background.	pre 2020
Openness to experience	Second factor of a principal component analysis of the items of the BIG 5-inventory.	2019
Conscientiousness	Third factor of a principal component analysis of the items of the BIG 5-inventory.	2019
Extraversion	First factor of a principal component analysis of the items of the BIG 5-inventory.	2019
Agreeableness	Fifth factor of a principal component analysis of the items of the BIG 5-inventory.	2019
Neuroticism	Fourth factor of a principal component analysis of the items of the BIG 5-inventory.	2019
Household size	Number of household members.	2019
Household net income	Monthly household net income in 2015 Euro. If information is missing, we imputed the information by plugging in the mean for each education x child presence x self-employment status-cell.	2019
Married	Indicator for being married.	2019
School child or younger	Indicator reflecting the presence of a child in school age or younger.	2020
Basic school leaving degree	Indicator for categories 0 "in school" to 1c "basic vocational education" according to the Comparative Analysis of Social Mobility in Industrial Nations (CASMIN)-scale.	Last available information in seven years pre 2020
Intermediate school leaving degree	Indicator for categories 2b "intermediate general qualification" to 2c_voc "vocational maturity certificate" according to the CASMIN-scale.	Last available information in seven years pre 2020
Tertiary school leaving degree	Indicator for categories 3a "lower tertiary education" or 3b "higher tertiary education" according to the CASMIN-scale.	Last available information in seven years pre 2020
Unemployment experience	Generated unemployment experience from "pgen.dta" of the SOEP v.35.	2018
NACE 2 code	Two-digit NACE Industry – Sector. Missing values, e.g. due to unemployment in 2019, are coded as separate category.	2019
Subject to regulation	Indicator reflecting whether self-employed individuals' business was subject to regulations to contain COVID-19, e.g. regulation of opening hours.	2020
Supply problems	Indicator reflecting whether self-employed individuals' business suffered from shortages of intermediate goods.	2020
Demand problems	Indicator reflecting whether the self-employed individuals' business suffered from cancellation of their services and goods, i.e. demand shortage.	2020

Note: Table 4.A.1 provides information on variables and their year of origin.

Table 4.A.2: Summary statistics

	(1) Self-employed	(2) Individuals	(3) Employees	(4) Individuals	(5) P-value of (1) -(3)
Income (gross) decrease	0.552	310	0.132	3,221	0.000
Working hour decrease	0.495	309	0.196	3,209	0.000
Remote work	0.457	311	0.390	3,222	0.021
Demographics:					
Age	53.791	311	47.034	3,222	0.000
	(11.154)		(10.533)		
Female	0.498	311	0.611	3,222	0.000
Migration background	0.164	311	0.205	3,222	0.086
Personality traits:					
Openness to experience	0.317	311	-0.032	3,222	0.000
	(1.010)			(0.975)	
Conscientiousness	0.099	311	0.076	3,222	0.664
	(0.928)			(0.919)	
Extraversion	0.092	311	0.015	3,222	0.196
	(0.967)			(1.019)	
Agreeableness	-0.005	311	-0.088	3,222	0.159
	(1.009)			(0.989)	
Neuroticism	-0.127	311	-0.051	3,222	0.188
	(0.954)			(0.973)	
Household context:					
Household size	2.617	311	2.815	3,222	0.017
	(1.427)			(1.386)	
Household net income (€)	4619.53	311	3826.88	3,222	0.000
	(4482.76)			(1970.61)	
Married	0.624	311	0.585	3,222	
School child or younger	0.354	311	0.468	3,222	0.000
Education (ref. basic)					
Intermediate	0.379	311	0.493	3,222	0.000
Tertiary	0.514	311	0.348	3,222	0.000
Unemployment experience	0.876	311	0.882	3,222	0.968
Revenue-reducing events in the wake of COVID-19:					
Subject to regulation	0.457	311			
Supply problems	0.122	311			
Demand problems	0.434	311			

Note: Table 4.A.2 displays mean and standard deviations, in parentheses, for self-employed and gainfully employed individuals.

Table 4.A.3: Summary statistics for self-employed individuals

	(1) Female	(2) Individuals	(3) Male	(4) Individuals	(5) P-value of (1) -(3)
Income (gross) decrease	0.632	155	0.471	155	0.004
Working hour decrease	0.536	153	0.455	156	0.156
Remote work	0.432	155	0.481	156	0.392
Demographics:					
Age	52.245	155	55.327	156	0.015
	(10.230)			(11.835)	
Female	1.000	155	0.000	156	.
Migration background	0.155	155	0.173	156	0.665
Personality traits:					
Openness to experience	0.232	155	0.403	156	0.135
	(1.015)			(1.001)	
Conscientiousness	0.144	155	0.055	156	0.397
	(0.939)			(0.918)	
Extraversion	0.235	155	-0.050	156	0.009
	(0.835)			(1.066)	
Agreeableness	0.199	155	-0.207	156	0.000
	(0.941)			(1.036)	
Neuroticism	0.042	155	-0.296	156	0.002
	(0.970)			(0.910)	
Household context:					
Household size	2.626	155	2.609	156	0.917
	(1.378)			(1.479)	
Household net income (€)	4374.67	155	4862.82	156	0.338
	(5021.36)			(3875.48)	
Married	0.613	155	0.635	156	
School child or younger	0.355	155	0.353	156	0.967
Education (ref. basic)					
Intermediate	0.413	155	0.346	156	0.226
Tertiary	0.484	155	0.545	156	0.283
Unemployment experience	0.868	155	0.883	156	0.965
Revenue-reducing events in the wake of COVID-19:					
Subject to regulation	0.561	155	0.353	156	0.000
Supply problems	0.110	155	0.135	156	0.504
Demand problems	0.458	155	0.410	156	0.397

Note: Table 4.A.3 displays mean and standard deviations, in parentheses, for self-employed individuals.

Table 4.A.4: Summary statistics for employees

	(1) Female	(2) Individuals	(3) Male	(4) Individuals	(5) P-value of (1) -(3)
Income (gross) decrease	0.123	1,969	0.146	1,252	0.063
Working hour decrease	0.205	1,959	0.182	1,250	0.121
Remote work	0.369	1,970	0.423	1,252	0.002
Demographics:					
Age	47.141	1,970	46.866	1,252	0.470
	(10.063)		(11.235)		
Female	1.000	1,970	0.000	1,252	.
Migration background	0.197	1,970	0.216	1,252	0.193
Personality traits:					
Openness to experience	-0.082	1,970	0.046	1,252	0.000
	(0.993)			(0.942)	
Conscientiousness	0.164	1,970	-0.063	1,252	0.000
	(0.904)		(0.925)		
Extraversion	0.110	1,970	-0.136	1,252	0.000
	(1.002)		(1.026)		
Agreeableness	0.036	1,970	-0.282	1,252	0.000
	(0.965)		(0.997)		
Neuroticism	0.100	1,970	-0.289	1,252	0.000
	(0.985)		(0.905)		
Household context:					
Household size	2.875	1,970	2.720	1,252	0.002
	(1.354)			(1.432)	
Household net income (€)	3763.45	1,970	3926.69	1,252	0.022
	(1936.66)			(2019.63)	
Married	0.580	1,970	0.593	1,252	
School child or younger	0.491	1,970	0.431	1,252	0.001
Education (ref. basic)					
Intermediate	0.535	1,970	0.427	1,252	0.000
Tertiary	0.327	1,970	0.382	1,252	0.001
Unemployment experience	0.985	1,970	0.719	1,252	0.004

Note: Table 4.A.3 displays mean and standard deviations, in parentheses, for employed individuals.

Table 4.A.5: Relevance of industry fixed effects in Table 4.1

		(1) Income	(2) Working hours	(3) Remote work
Model without industry fixed effects	Self-employed	0.434***	0.316***	0.014
		(0.029)	(0.030)	(0.031)
	R^2	0.12	0.07	0.21
Unrestricted model	Self-employed	0.421***	0.302***	0.021
		(0.031)	(0.031)	(0.032)
	R^2	0.23	0.13	0.31

Note: Table 4.A.5 displays the coefficient estimates and R-squared of the unrestricted models in Columns (2), (4), and (6) of Table 4.1 with and without the inclusion of industry fixed effects. Corresponding robust standard errors are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4.A.6: Comparison of the models for the likelihood of an income decrease for employed and self-employed individuals

	(1) Self-employed	(2) Employees	(3) P-value of (1)-(2)
Demographics:			
Gender: Female	0.081	0.014	0.285
	(0.073)	(0.013)	
Age	0.027	-0.004	0.057
	(0.019)	(0.005)	
Age squared	-0.000*	0.000	0.014
	(0.000)	(0.000)	
Migration background	0.064	0.041**	0.798
	(0.110)	(0.016)	
Big 5:			
Extraversion (2019)	0.011	-0.002	0.694
	(0.040)	(0.006)	
Conscientiousness (2019)	-0.031	-0.010	0.518
	(0.039)	(0.007)	
Openness to experience (2019)	0.066*	0.007	0.062
	(0.038)	(0.006)	
Neuroticism (2019)	-0.031	-0.005	0.389
	(0.036)	(0.006)	
Agreeableness (2019)	-0.040	0.000	0.173
	(0.035)	(0.006)	
Household context:			
HH Size (2019)	-0.061	0.009	0.037
	(0.039)	(0.007)	
Married	0.037	0.021	0.805
	(0.073)	(0.015)	
School child or younger	0.045	0.014	0.725
	(0.103)	(0.018)	
Log. of HH net income (2019/18)	-0.026	-0.028*	0.961
	(0.058)	(0.016)	
Education (ref. low):			
Intermediate education	-0.102	0.035*	0.198
	(0.125)	(0.019)	
High education	-0.149	0.018	0.135
	(0.132)	(0.021)	
Unemployment experience	-0.026**	0.003	0.007
	(0.012)	(0.003)	
Observations	310	3,221	
R^2	0.41	0.17	

Note: Table shows 4.A.6 separate models for employed and self-employed individuals. All models include state, week and industry fixed effects. The p-values are based on Chow test comparing coefficients after a seemingly unrelated regression. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.7: Comparison of the models for the likelihood of a working time decrease for employed and self-employed individuals

	(1) Self-employed	(2) Employees	(3) P-value of (1)-(2)
Demographics:			
Gender: Female	-0.051	0.026	0.220
	(0.073)	(0.016)	
Age	0.007	-0.008	0.408
	(0.020)	(0.006)	
Age squared	0.000	0.000	0.344
	(0.000)	(0.000)	
Migration background	0.120	0.031	0.295
	(0.099)	(0.019)	
Big 5:			
Extraversion (2019)	0.067*	0.005	0.052
	(0.037)	(0.007)	
Conscientiousness (2019)	-0.058	-0.014*	0.186
	(0.038)	(0.008)	
Openness to experience (2019)	0.051	0.002	0.113
	(0.036)	(0.008)	
Neuroticism (2019)	-0.003	-0.002	0.985
	(0.039)	(0.008)	
Agreeableness (2019)	-0.067*	-0.005	0.037
	(0.034)	(0.007)	
Household context:			
HH Size (2019)	-0.076**	0.016*	0.003
	(0.036)	(0.008)	
Married	-0.010	0.027	0.584
	(0.078)	(0.018)	
School child or younger	0.211**	-0.014	0.005
	(0.094)	(0.021)	
Log. of HH net income (2019/18)	0.100*	-0.044**	0.006
	(0.058)	(0.019)	
Education (ref. low):			
Intermediate education	0.074	0.016	0.551
	(0.114)	(0.023)	
High education	-0.026	-0.008	0.860
	(0.120)	(0.025)	
Unemployment experience	0.001	0.005*	0.668
	(0.010)	(0.003)	
Observations	309	3,209	
R^2	0.40	0.10	

Note: Table 4.A.7 shows separate models for employed and self-employed individuals. All models include state, week and industry fixed effects. The p-values are based on Chow test comparing coefficients after a seemingly unrelated regression. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.8: Comparison of the models for the likelihood of remote work for employed and self-employed individuals

	(1) Self-employed	(2) Employees	(3) P-value of (1)-(2)
Demographics:			
Gender: Female	-0.040	-0.009	0.612
	(0.069)	(0.018)	
Age	-0.042**	0.000	0.022
	(0.021)	(0.006)	
Age squared	0.000*	0.000	0.037
	(0.000)	(0.000)	
Migration background	-0.117	-0.020	0.191
	(0.085)	(0.019)	
Big 5:			
Extraversion (2019)	0.046	-0.007	0.093
	(0.037)	(0.008)	
Conscientiousness (2019)	0.033	-0.003	0.256
	(0.037)	(0.008)	
Openness to experience (2019)	0.058*	0.026***	0.272
	(0.034)	(0.008)	
Neuroticism (2019)	-0.013	-0.009	0.889
	(0.035)	(0.008)	
Agreeableness (2019)	-0.032	0.005	0.198
	(0.033)	(0.008)	
Household context:			
HH Size (2019)	0.092***	-0.019**	0.000
	(0.033)	(0.009)	
Married	0.026	-0.031*	0.356
	(0.071)	(0.019)	
School child or younger	-0.018	0.049**	0.436
	(0.101)	(0.023)	
Log. of HH net income (2019/18)	-0.146***	0.151***	0.000
	(0.052)	(0.020)	
Education (ref. low):			
Intermediate education	-0.108	0.069***	0.065
	(0.112)	(0.020)	
High education	0.057	0.283***	0.027
	(0.119)	(0.025)	
Unemployment experience	-0.013	-0.002	0.276
	(0.011)	(0.002)	
Observations	311	3,222	
R^2	0.47	0.34	

Note: Table 4.A.8 shows separate models for employed and self-employed individuals. All models include state, week and industry fixed effects. The p-values are based on Chow test comparing coefficients after a seemingly unrelated regression. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.9: Restricted and unrestricted model for difference of likelihood that income or working hours decreased or that the individual has transitioned into non-employment between employees and self-employed respondents, conditional on the employment status in 2019

	(1) Income	(2) Income	(3) Working hours	(4) Working hours	(5) Job loss	(6) Job loss
Self-employed	0.366***	0.364***	0.266***	0.267***	0.012	-0.007
	(0.031)	(0.033)	(0.031)	(0.033)	(0.009)	(0.018)
Demographics:						
Gender: Female		0.015		0.021		0.007
		(0.014)		(0.016)		(0.005)
Age		0.001		-0.003		-0.003*
		(0.005)		(0.006)		(0.002)
Age squared		0.000		0.000		0.000*
		(0.000)		(0.000)		(0.000)
Migration background		0.037**		0.042**		0.008
		(0.017)		(0.020)		(0.007)
Big 5:						
Extraversion (2019)		0.005		0.011		0.005**
		(0.006)		(0.007)		(0.002)
Conscientiousness (2019)		-0.008		-0.022***		-0.001
		(0.007)		(0.008)		(0.002)
Openness to experience (2019)		0.010		0.005		0.002
		(0.006)		(0.008)		(0.002)
Neuroticism (2019)		-0.003		0.002		0.002
		(0.006)		(0.008)		(0.002)
Agreeableness (2019)		0.001		-0.005		0.003
		(0.006)		(0.007)		(0.002)
Household context:						
HH Size (2019)		0.009		0.015*		0.001
		(0.008)		(0.009)		(0.003)
Married		0.015		0.014		0.005
		(0.016)		(0.018)		(0.006)
School child or younger		0.014		-0.005		0.000
		(0.019)		(0.021)		(0.007)
Log. of HH net income (2019/18)		-0.044***		-0.042**		-0.009
		(0.017)		(0.019)		(0.006)
Education (ref. low):						
Intermediate education		0.045**		0.023		-0.006
		(0.019)		(0.023)		(0.008)
High education		0.031		0.001		-0.001
		(0.022)		(0.025)		(0.009)
Unemployment experience		0.000		0.007*		0.004**
		(0.003)		(0.003)		(0.002)
Mean of outcome	0.168	0.168	0.219	0.219	0.017	0.017
Observations	3,348	3,348	3,334	3,334	3,661	3,661
R^2	0.08	0.22	0.04	0.13	0.01	0.05

Note: Table 4.A.9 displays models with and without controls for differences between self-employed and employees. All models include state and week fixed effects. Columns (1), (3) and (5) display results for the models without controls. Columns (2), (4) and (6) display results for the models with controls. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.10: Restricted and unrestricted model for likelihood that income and working hours decreased or individuals are working from home among employees

	(1) Income	(2) Income	(3) Working hours	(4) Working hours	(5) Remote work	(6) Remote work
Gender: Female	-0.022*	0.014	0.021	0.026	-0.048***	-0.009
	(0.012)	(0.013)	(0.014)	(0.016)	(0.018)	(0.018)
Demographics:						
Age		-0.004		-0.008		0.000
		(0.005)		(0.006)		(0.006)
Age squared		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
Migration background		0.041**		0.031		-0.020
		(0.016)		(0.019)		(0.019)
Big 5:						
Extraversion (2019)		-0.002		0.005		-0.007
		(0.006)		(0.007)		(0.008)
Conscientiousness (2019)		0.007		-0.014*		-0.003
		(0.006)		(0.008)		(0.008)
Openness to experience (2019)		-0.010		0.002		0.026***
		(0.007)		(0.008)		(0.008)
Neuroticism (2019)		-0.005		-0.002		-0.009
		(0.006)		(0.008)		(0.008)
Agreeableness (2019)		0.000		-0.005		0.005
		(0.006)		(0.007)		(0.008)
Household context:						
HH Size (2019)		0.009		0.016*		-0.019**
		(0.007)		(0.008)		(0.009)
Married		0.021		0.027		-0.031*
		(0.015)		(0.018)		(0.019)
School child or younger		0.014		-0.014		0.049**
		(0.018)		(0.021)		(0.023)
Log. of HH net income (2019/18)		-0.028*		-0.044**		0.151***
		(0.016)		(0.019)		(0.020)
Education (ref. low):						
Intermediate education		0.035*		0.016		0.069***
		(0.019)		(0.023)		(0.020)
High education		0.018		-0.008		0.283***
		(0.021)		(0.025)		(0.025)
Unemployment experience		0.003		0.005*		-0.002
		(0.003)		(0.003)		(0.002)
Mean of outcome	0.132	0.132	0.196	0.196	0.390	0.390
Observations	3,221	3,221	3,209	3,209	3,222	3,222
R^2	0.01	0.17	0.01	0.10	0.03	0.34

Note: Table 4.A.10 displays restricted and unrestricted models underlying the Gelbach decomposition. All models include state and week fixed effects. Columns (1), (3) and (5) display results for the restricted models. Columns (2), (4) and (6) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.11: Detailed results for the Gelbach decomposition of the gender gap among self-employed individuals

	(1) Income	(2) Working hours	(3) Remote work
Total change	0.093*	0.119**	0.022
	(0.049)	(0.049)	(0.051)
Demographics	0.031*	0.007	0.018
	(0.017)	(0.014)	(0.018)
NACE	0.092**	0.121***	0.000
	(0.045)	(0.043)	(0.041)
Big 5	-0.029	-0.010	-0.005
	(0.023)	(0.026)	(0.024)
Household context	-0.001	-0.003	0.016
	(0.012)	(0.014)	(0.016)
Unemployment experience	-0.002	0.000	0.001
	(0.007)	(0.001)	(0.004)
Education	0.001	0.004	-0.008
	(0.006)	(0.006)	(0.010)

Note: Table 4.A.11 displays the detailed results of the Gelbach decomposition of the gender gap among self-employed individuals. Columns (1), (2) and (3) display the results for the likelihood of an income decline, decline in working hours and working from home. The total change corresponds to the change in the gender gap between the restricted and the unrestricted models. The remaining rows show the contribution of the respective groups of covariates to the total change. Corresponding robust standard errors are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4.A.12: Detailed results for the Gelbach decomposition of the gender gap among employees

	(1) Income	(2) Working hours	(3) Remote work
Total change	-0.036	-0.005	-0.039***
	(0.009)	(0.010)	(0.013)
Demographics	-0.002	-0.002	0.000
	(0.001)	(0.002)	(0.002)
NACE	-0.036	-0.004	-0.021**
	(0.008)	(0.008)	(0.009)
Big 5	-0.005	-0.005	-0.007
	(0.004)	(0.005)	(0.005)
Household context	0.003*	0.002	-0.004
	(0.002)	(0.002)	(0.003)
Unemployment experience	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)
Education	0.003*	0.002	-0.006
	(0.001)	(0.002)	(0.004)

Note: Table 4.A.12 displays the detailed results of the Gelbach decomposition of the gender gap among employees. Columns (1), (2) and (3) display the results for the likelihood of an income decline, decline in working hours and working from home. The total change corresponds to the change in the gender gap between the restricted and the unrestricted models. The remaining rows show the contribution of the respective groups of covariates to the total change. Corresponding robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.13: The share of women and industry fixed effects for income losses

	Rank	NACE code	Description	Share female	FE estimate
	(1)	(2)	(3)	(4)	(5)
High share women	1	96	Other personal service activities	0.857	0.480**
					(0.236)
	2	88	Social work activities without accommodation	0.832	0.124
					(0.242)
	3	47	Retail trade, except of motor vehicles and motorcycles	0.818	0.775***
					(0.222)
	4	55	Accommodation	0.818	0.283
					(0.242)
	5	86	Human health activities	0.803	0.405*
					(0.208)
Low share women	1	49	Land transport and transport via pipelines	0.189	0.463
					(0.334)
	2	18	Printing and reproduction of recorded media	0.235	-0.425*
					(0.234)
	3	43	Specialized construction activities	0.273	0.093
					(0.249)
	4	62	Computer programming, consultancy and related activities	0.290	0.098
					(0.246)
	5	28	Manufacture of machinery and equipment n.e.c.	0.297	0.738***
					(0.218)

Note: Table 4.A.13 displays the share of women and the associated income loss fixed effects for the industries with the highest share and lowest share of women. For Table 4.A.13, we display only industries with at least ten observations. Column (1) displays the rank within each panel. Columns (2) and (3) display the two-digit NACE code and the description, respectively. Column (3) displays the share of women within each occupation in our full sample. Column (5) displays industry fixed-effect estimates, which stem from a regression of the likelihood of an income loss on state and week indicators as well as industry indicators, along with robust standard errors in parentheses The reference industry is "Crop and animal production, hunting and related service activities". * p<0.10, ** p<0.05, *** p<0.01

Table 4.A.14: Detailed results for the Gelbach decomposition of the gender gap in potential mechanisms among self-employed individuals

	(1) Restrictions	(2) Supply	(3) Demand
Total change	0.151***	0.030	0.059
	(0.049)	(0.033)	(0.050)
Demographics	0.017	0.022**	0.044**
	(0.015)	(0.011)	(0.018)
NACE	0.089*	0.022	0.021
	(0.046)	(0.028)	(0.043)
Big 5	0.046**	-0.012	-0.002
	(0.022)	(0.017)	(0.025)
Household context	0.000	-0.002	-0.004
	(0.010)	(0.007)	(0.011)
Unemployment experience	0.002	0.001	0.002
	(0.006)	(0.003)	(0.007)
Education	-0.003	-0.002	-0.001
	(0.005)	(0.006)	(0.005)

Note: Table 4.A.14 displays the detailed results of the Gelbach decomposition of the gender gap in potential mechanisms among self-employed individuals. Columns (1), (2) and (3) display the results for the likelihood of an income decline, decline in working hours and working from home. The total change corresponds to the change in the gender gap between the restricted and the unrestricted models. The remaining rows show the contribution of the respective groups of covariates to the total change. Corresponding robust standard errors are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4.A.15: Detailed results for the Gelbach decomposition of the gender gap in the likelihood of an income decline among self-employed individuals, including business-related events as an explanatory variable

	(1) Restrictions	(2) Supply	(3) Demand
Total change	0.103**	0.092*	0.089*
	(0.049)	(0.049)	(0.051)
NACE	0.071*	0.091**	0.083**
	(0.043)	(0.045)	(0.042)
Event	0.045**	-0.001	0.020
	(0.018)	(0.003)	(0.022)
Remainder	-0.013	0.001	-0.013

Note: Table 4.A.15 displays the detailed results of the Gelbach decomposition of the gender gap in the likelihood of an income decline among self-employed individuals. Columns (1), (2) and (3) display the results including and indicator whether respondents state their business has been affected by restrictions or policies, supply or demand shortages in the wake of the COVID-19 pandemic, respectively. The total change corresponds to the change in the gender gap between the restricted and the unrestricted models. The remaining characteristics are included in the group "Remainder". Corresponding robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

4.B Additional figures

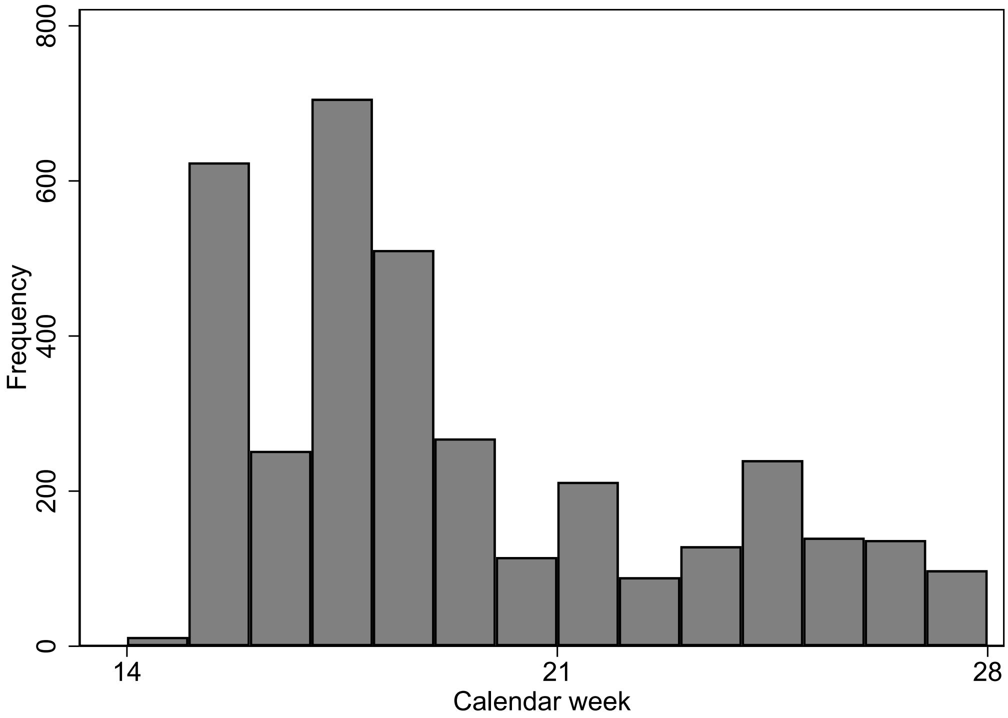
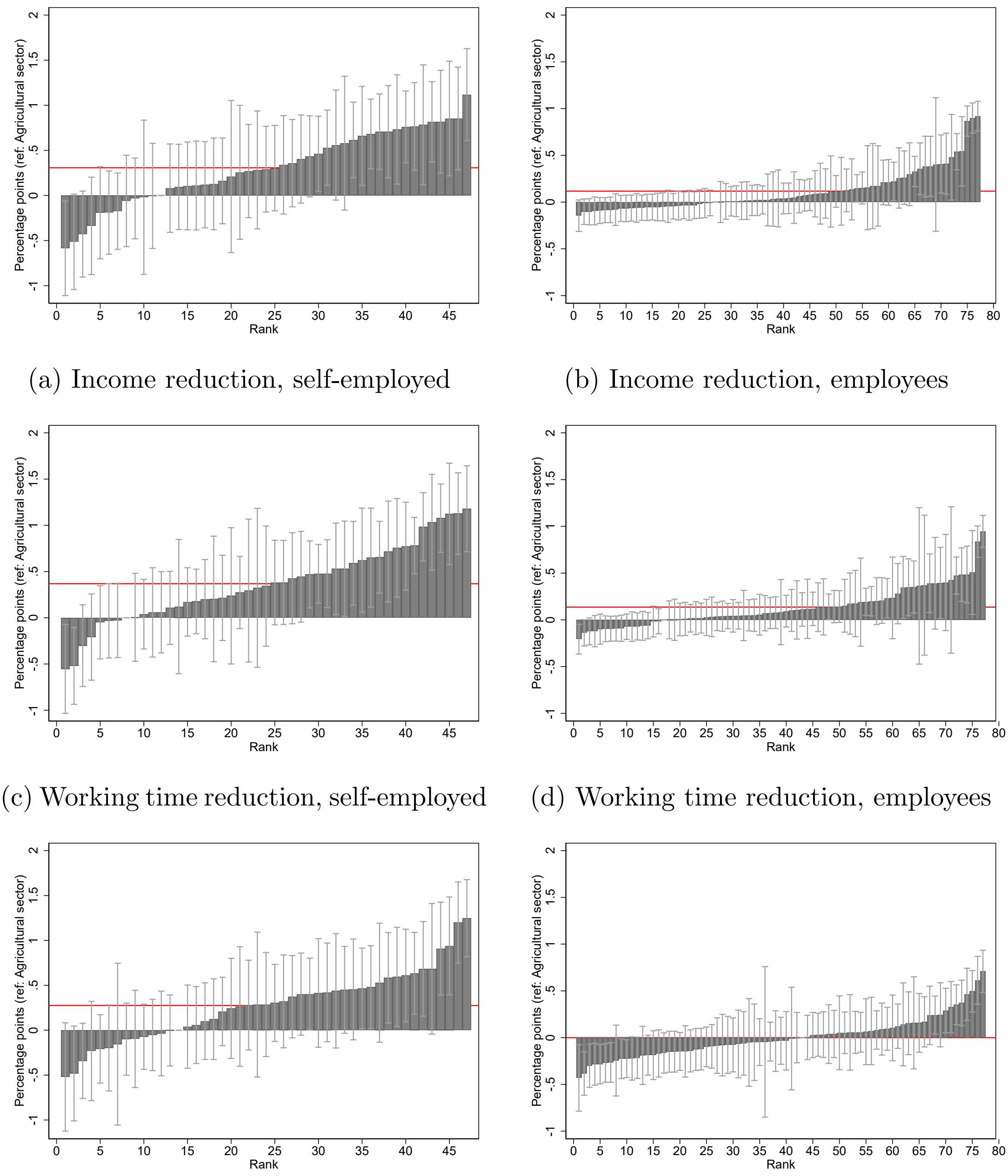


Figure 4.B.1: Distribution of observations over calendar weeks

Figure 4.B.2: Industry fixed effects for the self-employed and employees



(a) Income reduction, self-employed

(b) Income reduction, employees

(c) Working time reduction, self-employed

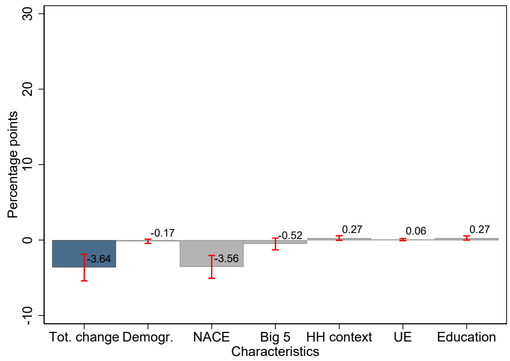
(d) Working time reduction, employees

(e) Home office, self-employed

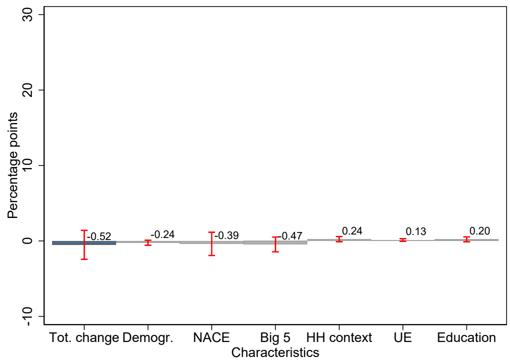
(f) Home office, employees

Note: Figures 4.B.2a to 4.B.2f display industry fixed effects and corresponding 95% confidence intervals from the regression results in Table 4.A.6 to 4.A.8. The horizontal line corresponds to the overall mean. Each rank corresponds to a specific industry (we use the two-digit NACE codes). Industries are ordered by the magnitude of their respective fixed effect. Since the sample size is smaller for the self-employed, there are fewer industries for which we have observations compared to employees, explaining the smaller number of ranks along the x-axis.

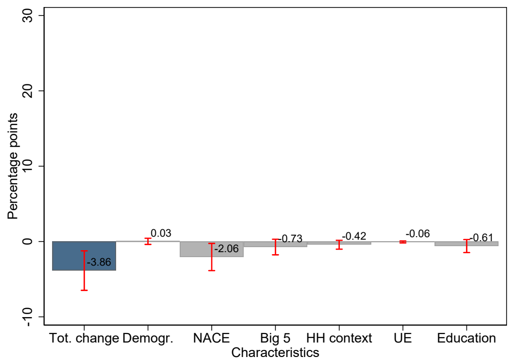
Figure 4.B.3: Gelbach decomposition of the gender gap in labor market outcomes among employees



(a) Likelihood of income decline



(b) Likelihood of decline in working time



(c) Likelihood of remote work

Note: Figures 4.B.3a to 4.B.3c display the Gelbach decomposition of the gender gap in the likelihood of an income, working time decline as well as the likelihood of working from home among employees. Red bars indicate 95% confidence intervals based on robust standard errors.

4.C Derivation of the Gelbach decomposition

Assume two sets of variables, X_1 and X_2, with k_1 and k_2 variables each.[30] The population linear relationship is given by:

$$Y = X_1\beta_1 + X_2\beta_2 + \epsilon \tag{4.1}$$

We label the components of the OLS estimator that correspond to the variables in X_1 and X_2, $\hat{\beta}_1^{full}$ and $\hat{\beta}_2$, respectively.

Thus, we obtain:

$$y = X_1\hat{\beta}_1^{full} + X_2\hat{\beta}_2 + \hat{\epsilon} \tag{4.2}$$

Now let us consider the coefficient on X_1 from a base specification that completely ignores the variables in X_2. We denote this estimator $\hat{\beta}_1^{base} = (X_1^\intercal X_1)^{-1}X_1^\intercal y$.

The Gelbach (2016) decomposition answers the question of how much of the change in X_1 coefficients can be attributed to different variables in X_2 as we move from the base specification that has no X_2 covariates to the full specification that includes both X_1 and all X_2 covariates. In the context of our analysis, X_1 would refer to a gender indicator, plus week and state fixed effects, and X_2 to the full set of control variables. The decomposition links the estimates of the base- and full-specification on X_1 through the following identity, which is obtained by pre-muliplying both sides of Equation 2 by $(X_1^\intercal X_1)^{-1}X_1^\intercal$ and using the orthogonality of the fitted residuals to the columns of X_1:

$$\hat{\beta}_1^{base} = \hat{\beta}_1^{full} + (X_1^\intercal X_1)^{-1}X_1^\intercal X_2\hat{\beta}_2 \tag{4.3}$$

Re-writing the above identity and defining the change in the coefficient on the gender dummy between the base and the full model as $\hat{\delta} \equiv \hat{\beta}_1^{base} - \hat{\beta}_1^{full}$, one obtains

[30] This exposition borrows heavily from the one given in Gelbach (2016).

$$\hat{\delta} \equiv \hat{\beta}_1^{base} - \hat{\beta}_1^{full} = (X_1^\intercal X_1)^{-1} X_1^\intercal X_2 \hat{\beta}_2, \tag{4.4}$$

which corresponds to the omitted variable bias formula.

Let X_{2k} be the column of observations on the k^{th} covariate in X_2 and let $\hat{\beta}_{2k}$ be the estimated coefficient on X_{2k} in the full specification, then

$$\hat{\delta} = \sum_{k=1}^{k_2} (X_1^\intercal X_1)^{-1} X_1^\intercal X_{2k} \hat{\beta}_{2k}, \tag{4.5}$$

since the omitted variables bias formula is linear in its k_2 components.

From there, the practical implementation of the decomposition follows naturally:

1. Estimate the full model to obtain $\hat{\beta}_2$.

2. Estimate the vector of coefficients on X_1 in a set of OLS regressions with each of the k_2 covariates X_{2k} as dependent variable. This yields $(X_1^\intercal X_1)^{-1} X_1^\intercal X_{2k}$.

3. Multiply $(X_1^\intercal X_1)^{-1} X_1^\intercal X_{2k}$ by $\hat{\beta}_{2k}$ to obtain $\hat{\delta}_k$, which is the component estimated to be due to each variable k.

The set of covariates we include in our Gelbach decomposition, i.e. X_2, are:

- Demographics: second-order polynomial in age, indicator for a migration background,

- NACE codes (2019): indicators for the two-digit NACE codes,

- Big 5 (2019): openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism,

- Household context (2019): household size, indicators for being married, presence of school children (or younger) in the household, the logarithm of household net income (2019/18) and

- unemployment experience (2018).

4.D Decomposing the gender gap in earnings losses due to the COVID-19 pandemic

In the following, we apply the Gelbach decomposition to the magnitude of the income losses and the reduction of working hours. The magnitude and directions of our estimates are consistent with our findings at the extensive margin. However, because of the reduction in the sample size, our estimations are not very precise. Consequently, the degree of statistical uncertainty is rather high. Note as well that relative changes are in relation to 2019 earnings and hours, respectively. Given that intra-year changes are frequent among the self-employed, the results should be interpreted with caution.

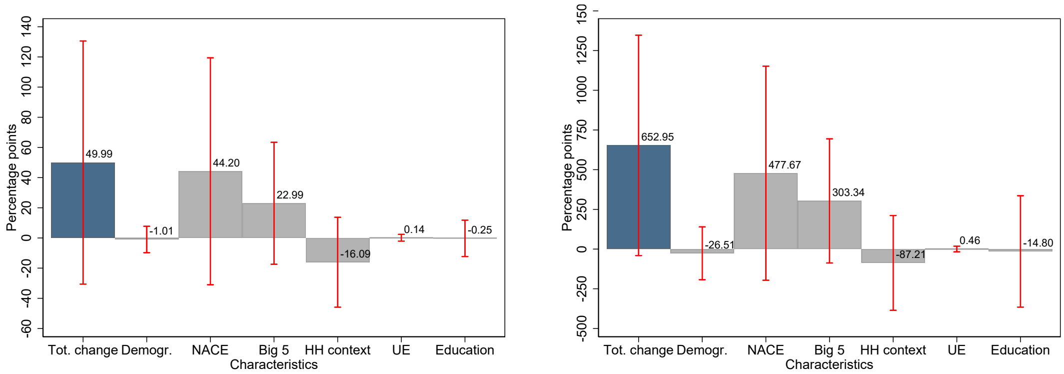
Table 4.D.1 displays the restricted and unrestricted model for the logarithm of monthly losses of gross earnings due to COVID-19 and relative monthly losses of gross earnings due to COVID-19. The result in column (1) of Table 4.D indicates that, on average, the losses of self-employed women are 70% smaller than for self-employed men. Once we include our full set of controls, this difference increases to approximately 120%. Figure 4.D.1a displays the Gelbach decomposition for the gender gap of the logarithm of monthly absolute earnings losses. Clearly, none of the components in Figure 4.D.1a are statistically different from zero. However, judging by the magnitude of the estimates, it appears that the estimates are economically significant. The total change in the gender gap when moving from the restricted to the unrestricted model amounts to approximately 50 percentage points. Consistent with our previous findings, the decomposition suggests that it is mainly the disproportionate representation of self-employed women in industries more strongly affected by the pandemic that explains the change in the estimate. The Gelbach decomposition suggests that this association accounts for about 44.2 percentage points or 88.4% of the total change of about 50 percentage points.

Table 4.D.1: Restricted and unrestricted model for relative and absolute monthly earnings losses among the self-employed

	(1) Absolute losses	(2) Absolut losses	(3) Relative losses	(4) Relative losses
Gender: Female	-0.702***	-1.202**	1.177	-5.354
	(0.252)	(0.488)	(1.127)	(3.503)
Demographics:				
Age		0.049		-0.114
		(0.149)		(1.142)
Age squared		0.000		0.004
		(0.002)		(0.010)
Migration background		0.170		1.041
		(0.552)		(2.587)
Big 5:				
Extraversion (2019)		0.248		0.989
		(0.242)		(1.079)
Conscientiousness (2019)		0.087		3.538**
		(0.186)		(1.358)
Openness to experience (2019)		-0.707**		-5.323**
		(0.300)		(2.015)
Neuroticism (2019)		-0.059		0.351
		(0.245)		(1.209)
Agreeableness (2019)		0.267		2.308*
		(0.186)		(1.183)
Household context:				
HH Size (2019)		-0.098		2.408
		(0.249)		(1.523)
Married		-0.712**		-2.181
		(0.333)		(2.731)
School child or younger		0.751		-1.197
		(0.674)		(3.865)
Log. of HH net income (2019/18)		0.361		-2.011
		(0.234)		(1.570)
Education (ref. low):				
Intermediate education		0.271		23.075***
		(0.728)		(7.424)
High education		0.382		25.763***
		(0.793)		(8.377)
Unemployment experience		-0.059		0.607
		(0.154)		(1.123)
Mean of outcome	7.279	7.279	1.542	1.542
Observations	104	104	81	81
R^2	0.23	0.78	0.19	0.88

Note: Table 4.D.1 displays restricted and unrestricted models underlying the Gelbach decomposition. All models include state and week fixed effects. Columns (1) and (3) display results for the restricted models. Columns (2) and (4) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Figure 4.D.1: Gelbach decomposition of the gender gap in monthly gross earnings losses



(a) Log of absolute earnings losses

(b) Relative earnings losses

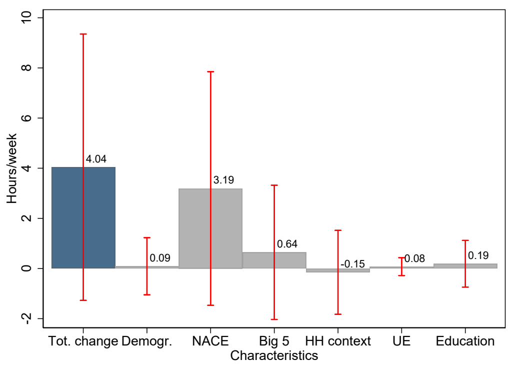
Note: Figures 4.D.1a and 4.D.1b display the Gelbach decomposition of the gender gap in the logarithm of absolute monthly gross earnings losses, and of the gender gap in relative monthly gross earnings losses among self-employed respondents. Red bars indicate 95% confidence intervals based on robust standard errors.

In columns (3) and (4) of Table 4.D.1, we display the results of the restricted and unrestricted models for relative earnings losses, where we divide the loss of monthly gross earnings by the level of monthly gross earnings in the previous year. The coefficient in the restricted model suggests that the relative losses were more than twice as large for self-employed women than for men. Clearly, this coefficient is not very precisely estimated. Thus, we are not able to reject the null of no difference by gender. Including all our controls causes the gender gap to decline by about 6.5, suggesting that women suffer five times less from relative earnings losses. Again, the coefficient is not very precisely estimated.

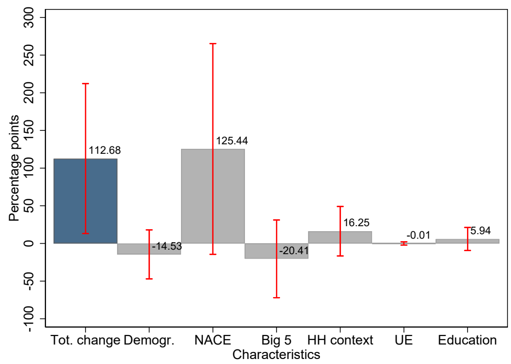
Figure 4.D.1b displays the results of the Gelbach decomposition for relative earnings losses. The inclusion of all our control variables causes the gender gap to decrease by about 653 percentage points. The total change is significant at the ten percent level of significance. Again, the Gelbach composition suggests that the largest fraction of this change is attributable to industry fixed effects. This indicates that self-employed women are working disproportionately in industries that are associated with larger relative earnings losses.

In Table 4.D.2, we display the restricted and unrestricted model for the logarithm of reduction of weekly working hours due to the COVID-19 pandemic, as well as for the relative reduction, where we again divide the decrease by the weekly working hours of the previous year. The raw gender gap in the reduction of working hours, displayed in column (1) in Table 4.D.2, amounts to -6%, which is not precisely estimated. If we include all our controls, the gender gap further decreases by about 19.3 percentage points. Therefore, the adjusted gender gap, displayed in column (2), amounts to -25.6%, which is not very precisely estimated. Turning to the Gelbach decomposition of absolute decreases in working hours, as depicted in Figure 4.D.2a, the estimates suggest that about 79% of the total change in the gender gap between the restricted and unrestricted model is again attributable to industry effects. While the changes are economically meaningful, the estimates are very imprecisely estimated and we are not able to reject the null hypotheses of no changes.

Figure 4.D.2: Gelbach decomposition of the gender gap in reductions of weekly working hours



(a) Reduction of weekly working hours



(b) Rel. reduction in weekly working hours

Note: Figures 4.D.2a and 4.D.2b display the Gelbach decomposition of the gender gap in the logarithm of reductions in weekly working hours, and of the gender gap in relative weekly working hours reductions among self-employed respondents. Red bars indicate 95% confidence intervals based on robust standard errors.

Lastly, the restricted and unrestricted model for the relative changes in weekly working hours are displayed in columns (3) and (4) of Table 4.D.2. The raw gender gap amounts to a very imprecisely estimated 15.5 percentage points. Thus, the relative change in weekly working hours of self-employed women are 15.5 percentage points higher. If we include our complete set of controls, the gender gap declines by about 112.7 percentage points. Thus, the adjusted gender gap in relative reductions of working hours reverses in sign and amounts to -97.2 percentage points, which again is very imprecisely estimated. Turning to the Gelbach decomposition, depicted in Figure

Table 4.D.2: Restricted and unrestricted model for relative and absolute change in weekly working hours due to the COVID-19 pandemic

	(1) Absolute change	(2) Absolute change	(3) Relative change	(4) Relative change
Gender: Female	-0.063	-0.256	0.155	-0.972
	(0.147)	(0.195)	(0.224)	(0.652)
Demographics:				
Age		0.174**		0.143
		(0.066)		(0.188)
Age squared		-0.002**		-0.001
		(0.001)		(0.002)
Migration background		0.329		-0.548
		(0.204)		(0.421)
Big 5:				
Extraversion (2019)		0.006		0.289
		(0.102)		(0.191)
Conscientiousness (2019)		0.160**		-0.003
		(0.077)		(0.209)
Openness to experience (2019)		0.105		0.175
		(0.122)		(0.185)
Neuroticism (2019)		0.080		-0.221
		(0.112)		(0.236)
Agreeableness (2019)		-0.053		-0.326
		(0.089)		(0.434)
Household context:				
HH Size (2019)		0.004		-0.347
		(0.086)		(0.207)
Married		-0.498**		-0.034
		(0.203)		(0.400)
School child		0.445		1.812*
		(0.269)		(0.928)
Log. of HH net income (2019/18)		-0.020		0.074
		(0.145)		(0.258)
Education (ref. low):				
Intermediate education		-0.068		-1.093
		(0.364)		(0.867)
High education		-0.185		-0.897
		(0.352)		(0.879)
Unemployment experience		-0.018		-0.001
		(0.017)		(0.083)
Mean of outcome	18.068	18.068	0.780	0.780
Observations	148	148	122	122
R^2	0.17	0.64	0.15	0.64

Note: Table 4.D.2 displays restricted and unrestricted models underlying the Gelbach decomposition. All models include state and week fixed effects. Columns (1) and (3) display results for the restricted models. Columns (2) and (4) display results for the unrestricted models. The unrestricted models also include NACE 2 fixed effects. Standard errors are robust and in parentheses. * p<0.10, ** p<0.05, *** p<0.01

4.D.2b, we observe that the total change of 112.7 percentage points is statistically significant. Finally, the contribution of industry fixed effects to the total change amounts to 125.4 and is significant at the ten percent level of significance.

Summary and overall conclusion

Socio-economic differences in health or consequences due to public health shocks are present in all modern societies. In addition, there is overwhelming social consensus that these health differences are unfair when caused by differences in circumstances such as family backgrounds, migration status or gender. In many countries, public health efforts are directed at eradicating differences in health or socio-economic differences caused by pandemic shocks that are rooted in socio-economic differences. Developing targeted measures to meet this aim requires compelling evidence on the magnitude and causes of socio-economic differences in health or public health shocks. This Ph.D. thesis makes key contributions to measuring and understanding the health differences that are rooted in these circumstances and to studying the circumstances' impacts in the wake of a public health crisis. In the following, I briefly summarize the individual chapters and their implications for public policy.

Chapter 1 provides the first quantification of intergenerational mobility in permanent health in Germany. Using data from the Socio-Economic Panel (SOEP), providing more than 25 years of rich health information, Chapter 1 presents rank-rank regressions of the children's rank on the parental rank in the permanent health distribution. The results indicate that the rank-rank slope corresponds to 0.232 and the estimates of upward and downward mobility are 44.43 and 56.54, respectively. The estimates of the rank-rank slope for Germany are of similar magnitude to comparable estimates of intergenerational mobility in permanent income. Further, compared to the United States and Denmark, Germany ranks just in the middle, mirroring the country ranking in intergenerational income mobility.

If permanent health is anchored in permanent income, the results indicate that a one percentile increase in the permanent health distribution is associated with a 0.8 to 1.4% increase in permanent income. However, strong nonlinearities in this association

are present at the lower end of the permanent health distribution.
Finally, the results of Chapter 1 indicate that a more favorable SES of the parents is associated with higher upward mobility in permanent health. This is an important difference from studies using U.S. data, which show that better parental SES is associated with better health of children throughout the parental permanent health distribution. Chapter 1 concludes by arguing that intergenerational mobility informs the reader about how equitable a society is.

Chapter 2 presents the first estimates of the effect of mothers' education on their children's mental health in adulthood. This is an important question since mental health is one of the leading sources for the high costs of non-communicable diseases.
To derive consistent estimates of the effect of maternal education on children's mental health in adulthood, we use exogenous variation in maternal schooling caused by a compulsory schooling law reform that extended the number of years of compulsory schooling from eight to nine. This analysis relies on the SOEP data. The data on children's mental health stems from the Mental Component Summary (MCS) score, a summary score for mental health. We also provide evidence on the dimension of children's physical health as captured by the Physical Component Summary (PCS) score. The PCS score is the equivalent to the MCS score for the dimension of physical health. Both measures are derived from a principal component analysis of the 12 items of the Short Form-12 questionnaire (SF-12).
The results in Chapter 2 indicate the absence of any effect of the number of years of maternal schooling on children's mental health in adulthood. However, previous results on the effects of the number of years of maternal schooling on children's physical health are replicated. Further results indicate that it is mainly the children's physical functioning that is positively affected.
Further, while the reduced form effect of maternal years of schooling reveals no effect, this does not preclude the existence of mediators. We therefore test potential mediators and find evidence that the number of friends, a frequently used proxy for social capital, is a mediator of the number of years of maternal schooling on children's mental health in adulthood. However, the implied overall effect is only very small, consistent with a zero overall effect.

Chapter 3 adds to the literature on health differences between migrants and the resident population by showing the effects of hate crime on refugees' mental health. This is of particular relevance in light of the stark increase in the number of refugees in Germany over the last decade and the increase in the number of hate crimes over the same period.
Consistent estimates are derived by employing a regression discontinuity design in time and the IAB-BAMF-SOEP Survey of Refugees, a special survey on refugees in Germany. The mental health outcomes are the MCS score and the Patient Health Questionnaire 4 (PHQ-4) score, the latter being a measure of the frequency of symptoms of depression and anxiety.
The results show that hate crimes reduce the MCS and PHQ-4 score by about 37 and 28% of a standard deviation, respectively. Further, the results presented in Chapter 3 show that country-specific human capital, such as language proficiency and number of German friends, moderates the effect considerably. This points to the importance of information acquisition as an ability that allows individuals to bring subjective perceptions into line with the true probability of becoming a victim of hate crime.

Chapter 4 shows how a public health crisis such as the COVID-19 pandemic leads to differences in economic outcomes between men and women. The COVID-19 pandemic is probably the greatest challenge to modern societies since World War II. It has unleashed severe economic crises in countries worldwide and led to the development of policies aimed at reducing the spread of the virus. Chapter 4 shows that the COVID-19 pandemic caused self-employed women's likelihood of an income reduction to be 35% higher than for self-employed men. Further, Chapter 4 shows that the effect is largely driven by the disproportionate representation of women in sectors most severely affected by the COVID-19 pandemic. These gender differences emerged mainly because the sectors in which women are more likely to work are more likely to be affected by government regulations aimed at combatting the pandemic.
In the following, I briefly summarize how each chapter informs public policy. Chapter 1 and 2 focus on the family background as potential circumstances driving health differences. Evidence on this is particular compelling since policy measures designed to compensate for health differences that are rooted in childhood are often very expensive.

If these resources could be shifted to earlier interventions, this could provide scope for efficiency gains. Chapter 1 contributes to this by providing important evidence on the persistence in permanent health across generations and on how differences in health translate into differences in permanent income. Moreover, while no causal claims are possible based on this evidence, we find that a favorable SES is most often associated with higher upward mobility. Holding mobility at all other ranks fixed, this could provide a feasible way to achieve Pareto improvements in health.

Moreover, our finding that the number of years of maternal schooling at the lower end of the distribution does not have any effect on children's mental health could inform public health efforts. This finding does not, however, rule out any effect of maternal education on children's mental health, given that the compulsory schooling law reform had no legal consequences in terms of access to different vocational or university education. Research on this relationship at different educational margins would be a fruitful venue for future research.

Our finding of the large effect of hate crimes on refugees' mental health, and the potential impact on the integration and long-term success of refuges and their children should be of utmost importance to policy makers. Evidence has shown that hate crimes limit refugees' potential to assimilate successfully and hinder the host country's economic growth. Chapter 3 should therefore motivate policy makers to shift resources into fostering a welcoming atmosphere for refugees and into mental health resources for refugees.

Chapter 4 shows how an otherwise well-intended and urgently needed policy measure to stop the spread of a communicable disease can have differential economic effects on women and men. Policy makers should consider these differential impacts and aim at designing compensation schemes that are universal, but proportional, to eradicate the existing differences. Failing to do so bears the risk of underutilizing the economic potential of the self-employed, and self-employed women in particular, who are an important source of innovation, and hence, long-term growth.

German summary

Sozioökonomische Unterschiede beim Gesundheitszustand oder in den Folgen von Gesundheitsschocks zeigen sich in allen modernen Gesellschaften. Es besteht gesellschaftlicher Konsens darin, dass diese gesundheitlichen Unterschiede ungerecht sind, insofern sie durch unterschiedliche Lebensumstände, wie den familiären Hintergrund, Migrationsstatus oder Geschlecht, verursacht werden. In vielen Ländern zielen die Bemühungen von Politikmaßnahmen darauf ab, gesundheitliche Unterschiede, die durch unterschiedliche sozioökonomische Umstände verursacht werden, sowie Unterschiede, die auf die Ausbreitung von Krankheiten zurück zu führen sind, zu beseitigen. Die Entwicklung passgenauer Maßnahmen zur Erreichung dieser Ziele erfordert Erkenntnisse über die diesen Phänomenen zu Grunde liegenden Prozesse. Diese Dissertation leistet wichtige Beiträge zur Messung und zum Verständnis ebendieser Prozesse. Im Folgenden fasse ich die einzelnen Kapitel der Dissertation sowie deren Implikationen für die Gestaltung von Politikmaßnahmen kurz zusammen.

Kapitel 1 beschreibt die erste Quantifizierung der intergenerationalen Mobilität in der permanenten Gesundheit in Deutschland. Unter Verwendung des Sozio-ökonomischen Panels (SOEP), das über mehr als 25 Jahre umfassender Gesundheitsinformationen zur Verfügung stellt, werden in Kapitel 1 Rangordnungsregressionen des Perzentil-Rangs der Kinder auf den elterlichen Perzentil-Rang in der Verteilung der permanenten Gesundheit vorgestellt. Die Ergebnisse zeigen, dass die Rang-Rang-Steigung 0,232 entspricht und die Schätzungen der Aufwärts- und Abwärtsmobilität 44,43 bzw. 56,54 betragen. Die Schätzungen der Rang-Rang-Steigung für Deutschland liegen in einer ähnlichen Größenordnung wie vergleichbare Schätzungen der intergenerationalen Mobilität für das permanente Einkommen. Darüber hinaus liegt Deutschland im Vergleich zu den USA und Dänemark bezüglich der intergenerationalen Mobilität im Mittelfeld, was die Rangfolge der Länder bei der intergenerationalen Einkommensmobilität widerspiegelt. Die Ergebnisse zeigen auch, dass ein Anstieg von einem Perzentil-Rang in der Verteilung der permanenten Gesundheit mit einem Anstieg des permanenten Einkommens um 0,8 bis 1,4% verbunden ist. Am unteren Ende der Verteilung der dauerhaften Gesundheit ist dieser Zusammenhang jedoch stark nichtlinear. Das heißt, Veränderungen in der Verteilung der permanenten Gesundheit am unteren Ende der Verteilung sind hier besonders relevant für das permanente Einkommen.

Darüber hinaus deuten die Ergebnisse von Kapitel 1 darauf hin, dass ein höherer sozioökonomischer Status der Eltern mit einer höheren Aufwärtsmobilität bei der permanenten Gesundheit verbunden ist. Dies ist ein wichtiger Unterschied zu Studien aus den Vereinigten Staaten, die zeigen, dass ein besserer elterlicher sozioökonomischer Status mit einer besseren Gesundheit der Kinder über die gesamte elterliche Verteilung der dauerhaften Gesundheit verbunden ist. Kapitel 1 schließt mit dem Argument, dass die intergenerationale Mobilität in Gesundheit Aufschluss darüber gibt, wie gerecht eine Gesellschaft ist.

Kapitel 2 präsentiert die ersten Schätzungen des Effekts der Bildung der Mütter auf die psychische Gesundheit ihrer Kinder im Erwachsenenalter. Dies ist eine wichtige Frage, da psychische Erkrankungen eine der Hauptursachen für die hohen Kosten von nicht übertragbaren Krankheiten sind.

Um konsistente Schätzungen des Effekts der mütterlichen Bildung auf die psychische Gesundheit der Kinder im Erwachsenenalter zu erzielen, verwenden wir exogene Variation in der mütterlichen Schulbildung, die sich durch eine Reform des Schulpflichtgesetzes ergibt, in deren Rahmen die Anzahl der Pflichtschuljahre von acht auf neun erhöht wurde. Diese Analyse stützt sich auf die Daten des SOEP. Die Daten zur psychischen Gesundheit der Kinder beruhen auf dem Mental Component Summary (MCS) Score, einem Index für die allgemeine psychische Gesundheit. Wir liefern auch Erkenntnisse über die Dimension der körperlichen Gesundheit der Kinder, die durch den Physical Component Summary (PCS) Score erfasst wird. Der PCS Score ist das Äquivalent zum MCS Score für die Dimension der physischen Gesundheit. Beide Maße werden aus einer Hauptkomponentenanalyse der 12 Items des Short Form-12 (SF-12)-Fragebogens abgeleitet.

Die Ergebnisse in Kapitel 2 deuten darauf hin, dass die Anzahl der Jahre der Schulbildung der Mutter keinen Einfluss auf die psychische Gesundheit der Kinder im Erwachsenenalter hat. Allerdings werden frühere Ergebnisse zur Anzahl der Jahre mütterlicher Schulbildung auf die physische Gesundheit der Kinder repliziert. Weitergehende Analysen deuten darauf hin, dass vor allem die körperlichen Funktionen der Kinder positiv beeinflusst werden. Dieses Ergebnis konnte bisher in der ökonomischen Literatur nicht gezeigt werden.

Zwar deuten die Schätzungen der mütterlichen Schuljahre auf die psychische Gesundheit der Kinder im Erwachsenenalter auf die Abwesenheit eines Effekts hin, dies schließt jedoch die Existenz von Mediatoren des betrachteten Zusammenhangs nicht aus. Wir testen daher potenzielle Mediatoren und finden Hinweise darauf, dass die Anzahl der Freunde, ein häufig verwendetes Maß für soziales Kapital, ein Mediator des Zusammenhangs zwischen der Anzahl der mütterlichen Schuljahre und der psychischen Gesundheit der Kinder im Erwachsenenalter ist. Der implizierte Gesamteffekt des Mediators ist jedoch nur sehr klein, was mit einem Gesamteffekt von Null konsistent ist.

Kapitel 3 ergänzt die Literatur zu gesundheitlichen Unterschieden zwischen Migranten und der einheimischen Bevölkerung, indem es die Auswirkungen von Hasskriminalität auf die psychische Gesundheit von Geflüchteten aufzeigt. Dies ist von besonderer Relevanz, sind doch die Anzahl der Geflüchteten und die Häufigkeit von Hasskriminalität im gleichen Zeitraum sprunghaft angestiegen.

Konsistente Schätzungen werden durch eine Regressionsdiskontinuitätsanalyse im Zeitverlauf und der IAB-BAMF-SOEP-Befragung von Geflüchteten, einer Sondererhebung zu Geflüchteten in Deutschland, erzielt. Die Maße für die psychische Gesundheit in dieser Studie sind der MCS-Score und der Patient Health Questionnaire-4 (PHQ-4) Score. Letzterer ist ein Maß für die Häufigkeit von Depressions- und Angstsymptomen.

Die Ergebnisse zeigen, dass Hasskriminalität den MCS und PHQ-4 Score um etwa 37 bzw. 28% einer Standardabweichung reduzieren. Weiterhin zeigen die in Kapitel 3 vorgestellten Ergebnisse, dass länderspezifisches Humankapital, wie Sprachkenntnisse und Anzahl der deutschen Freunde, den Effekt moderiert. Dies weist auf die Bedeutung der Möglichkeit zur Informationsbeschaffung hin, die Geflüchteten hilft ihre subjektive Wahrnehmung mit der tatsächlichen Wahrscheinlichkeit, Opfer von Hasskriminalität zu werden, in Einklang zu bringen.

Kapitel 4 zeigt, wie sich eine öffentliche Gesundheitskrise, wie die COVID-19-Pandemie, auf unterschiedliche Weise auf die wirtschaftlichen Ergebnisse von Männern und Frauen auswirken kann. Die COVID-19-Pandemie ist wahrscheinlich die größte Herausforderung für moderne Gesellschaften seit dem Zweiten Weltkrieg. Sie hat in Ländern auf der ganzen Welt schwere Wirtschaftskrisen ausgelöst und zur Entwicklung von Maßnahmen geführt, die darauf abzielen, die Ausbreitung des Virus zu reduzieren. Kapitel 4 zeigt, dass die COVID-19-Pandemie dazu führte, dass die Wahrscheinlichkeit einer Einkommensminderung bei selbständigen Frauen um rund 35% höher war als bei selbständigen Männern. Des Weiteren zeigt Kapitel 4, dass dieser Effekt größtenteils auf die überproportionale Selektion von Frauen in die von der COVID-19-Pandemie am stärksten betroffenen Branchen zurückzuführen ist. Diese geschlechtsspezifischen Unterschiede sind auch deshalb entstanden, weil die Sektoren, in denen Frauen mit größerer Wahrscheinlichkeit arbeiten, stärker von staatlichen Regelungen zur Bekämpfung der Pandemie betroffen sind.

Im Folgenden fasse ich kurz zusammen, wie die Ergebnisse der jeweiligen Kapitel das Design von verschieden Politikmaßnahmen beeinflussen können. Kapitel 1 und 2 konzentrieren sich auf den familiären Hintergrund als mögliche Ursache für gesundheitliche Unterschiede. Die dort zu Tage gebrachten Erkenntnisse sind besonders relevant, da politische Maßnahmen zum Ausgleich von Gesundheitsunterschieden, die in der Kindheit wurzeln, oft mit großen Kosten assoziiert sind. Wenn diese Ressourcen auf im Lebenszyklus frühe Interventionen verlagert werden könnten, könnte dies Spielraum für Effizienzgewinne bieten. Kapitel 1 trägt hierzu ebenfalls bei, indem es wichtige Erkenntnisse über die Persistenz der dauerhaften Gesundheit über Generationen hinweg liefert und darüber, wie sich Unterschiede in der permanenten Gesundheit in Unterschiede im permanenten Einkommen niederschlagen. Auch wenn auf Basis dieser Evidenz keine kausalen Behauptungen möglich sind, finden wir, dass ein günstiger sozioökonomischer Hintergrund der Eltern häufig mit einer höheren Aufwärtsmobilität verbunden ist. Hält man die Mobilität auf allen anderen Perzentil-Rängen konstant, könnte dies ein gangbarer Weg sein, um Pareto-Verbesserungen in der Gesundheit zu erreichen. Darüber hinaus könnte unser Befund, dass die Anzahl der Schuljahre der Mütter am unteren Ende der Bildungsverteilung keinen Einfluss auf die psychische

Gesundheit der Kinder hat, wichtig für die Bemühungen des öffentlichen Gesundheitswesens sein, den sozioökonomischen Gradienten in psychischer Gesundheit zu verringern. Dieser Befund schließt jedoch einen Effekt der mütterlichen Bildung auf die psychische Gesundheit der Kinder nicht aus, da die Reform des Schulpflichtgesetzes keine rechtlichen Konsequenzen in Bezug auf den Zugang zu verschiedenen Berufs- oder Hochschulausbildungen hatte. Die Erforschung dieses Zusammenhangs an unterschiedlichen Bildungsrändern wäre eine vielversprechende Möglichkeit für zukünftige Forschung.

Unsere Erkenntnisse über die Auswirkungen von Hassverbrechen auf die psychische Gesundheit von Geflüchteten und die potenziellen Auswirkungen auf die Integration und den langfristigen Erfolg von Geflüchteten und ihren Kindern sollte ebenfalls für politische Entscheidungsträger von größter Bedeutung sein. Bisherige Forschungsergebnisse legen nahe, dass Hassverbrechen die Integration von Geflüchteten behindern und diese daher nicht entsprechend ihres eigentlichen Potenzials zum Wirtschaftswachstum des Aufnahmelandes beitragen können. Kapitel 3 sollte daher die politischen Entscheidungsträger motivieren, Ressourcen in die Förderung einer Willkommensatmosphäre für Geflüchtete sowie ihrer psychischen Gesundheit zu investieren.

Kapitel 4 zeigt, wie eine dringend notwendige politische Maßnahme zur Verhinderung der Ausbreitung einer übertragbaren Krankheit unterschiedliche wirtschaftliche Auswirkungen auf Frauen und Männer haben kann. Politische Entscheidungsträger sollten diese unterschiedlichen Auswirkungen berücksichtigen und darauf abzielen, Ausgleichsregelungen zu treffen, die universell im Anspruch, aber proportional zur Betroffenheit sind, um die entstandenen Unterschiede zu beseitigen. Geschieht dies nicht, besteht die Gefahr, dass das wirtschaftliche Potenzial der Selbstständigen, und insbesondere der selbstständigen Frauen, die eine wichtige Quelle für Innovationen und damit für langfristiges Wachstum sind, nicht ausreichend genutzt wird.

Eidesstattliche Erklärung

und Einverständniserklärung nach §6 Abs. 2 Nr. 5, 6 und 7 der Promotionsordnung der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Potsdam vom 10.07.2013. Hiermit versichere ich an Eides statt, dass

- meine hinsichtlich der früheren Teilnahme an Promotionsverfahren gemachten Angaben richtig sind;
- die eingereichte Arbeit oder wesentliche Teile derselben in keinem anderen Verfahren zur Erlangung eines akademischen Grades vorgelegt worden sind;
- bei der Anfertigung der Dissertation die Grundsätze zur Sicherung guter wissenschaftlicher Praxis der DFG eingehalten wurden, die Dissertation selbständig und ohne fremde Hilfe verfasst wurde, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt worden sind und die den benutzten Werken wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht wurden.

- Ich erkläre hiermit, dass ich folgende Hilfsmittel genutzt habe:
  - Regressionen und Statistiken: Stata
  - Digitalisierung von Daten: Python
  - Georeferenzierung: R
  - Schriftsatz und Formatierungen: LaTeX, Overleaf, Git

Auf dieser Grundlage habe ich die Arbeit selbständig verfasst.

Einer Überprüfung der eingereichten Dissertationsschrift bzw. der an deren Stelle eingereichten Schriften mittels einer Plagiatssoftware stimme ich zu.

Ort & Datum

Unterschrift

Bibliography

Adams-Prassl, A., T. Boneva, M. Golin, and C. Rauh (2020). Inequality in the impact of the coronavirus shock: Evidence from real time surveys. *Journal of Public Economics 189*, online first.

Adermon, A., M. Lindahl, and M. Palme (2019). Dynastic Human Capital, Inequality and Intergenerational Mobility. IZA Discussion Papers 12300, Institute of Labor Economics (IZA).

Adhvaryu, A., J. Fenske, N. Kala, and A. Nyshadham (2015). Fetal origins of mental health: evidence from Africa. (2015-15).

Adhvaryu, A., J. Fenske, and A. Nyshadham (2019). Early life circumstance and adult mental health. *Journal of Political Economy 127*(4), 1516–1549.

Ahlburg, D. (1998). Intergenerational transmission of health. *The American Economic Review 88*(2), 265–270.

Akay, A., O. Bargain, and A. Elsayed (2020). Global terror, well-being and political attitudes. *European Economic Review 123*, online first.

Alesina, A., S. Hohmann, S. Michalopoulos, and E. Papaioannou (2021). Intergenerational mobility in Africa. *Econometrica 89*(1), 1–35.

Alidou, S. and M. Verpoorten (2019). Family size and schooling in sub-saharan Africa: testing the quantity-quality trade-off. *Journal of Population Economics 32*(4), 1353–1399.

Alipour, J.-V., O. Falck, and S. Schüller (2020). Germany's capacities to work from home. CESifo Working Paper 8227, Center for Economic Studies and ifo Institute (CESifo).

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–172.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature 56*(4), 1360–1446.

Alon, T. M., M. Doepke, J. Olmstead-Rumsey, and M. Tertilt (2020). The impact of COVID-19 on gender equality. Working Paper 26947, National Bureau of Economic Research, Inc.

Amin, V., C. A. Flores, and A. Flores-Lagunes (2020). The impact of BMI on mental health: Further evidence from genetic markers. *Economics & Human Biology 38*, online first.

Andersen, C. (2019). Intergenerational health mobility: Evidence from Danish registers. Economics Working Papers 2019-04, Department of Economics and Business Economics, Aarhus University.

Andersen, H. H., A. Mühlbacher, and M. Nübling (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF-12v2. *Schmollers Jahrbuch 127*, 171–182.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D., V. Lavy, and A. Schlosser (2005). New evidence on the causal link between the quantity and quality of children. NBER Working Papers 11835, National Bureau of Economic Research, Inc.

Antecol, H. and K. Bedard (2006). Unhealthy assimilation: Why do immigrants converge to American health status levels? *Demography 43*(2), 337–360.

Antman, F. M., B. Duncan, and S. J. Trejo (2020). Ethnic attrition, assimilation, and the measured health outcomes of Mexican Americans. *Journal of Population Economics 33*(4), 1499–1522.

Apouey, B. and P.-Y. Geoffard (2013). Family income and child health in the UK. *Journal of Health Economics 32*(4), 715 – 727.

Audretsch, D. B., A. S. Kritikos, and A. Schiersch (2020). Microfirms and innovation in the service sector. *Small Business Economics 55*(4), 997–1018.

Avendano, M., A. de Coulon, and V. Nafilyan (2020). Does longer compulsory schooling affect mental health? Evidence from a British reform. *Journal of Public Economics 183*, online first.

BAMF (2015). Verfahrensregelung zur Aussetzung des Dublinverfahrens für syrische Staatsangehörige. Interner Erlass Az. 411 - 93605/Syrien/2015, BAMF.

BAMF (2016). Aktuelle Zahlen zu Asyl, Ausgabe: Dezember 2016. Tabellen, Diagramme, Erläuterungen. Technical report, Federal Office for Migration and Refugees.

Banks, J. and F. Mazzonna (2012). The effect of education on old age cognitive abilities: Evidence from a regression discontinuity design. *The Economic Journal 122*(560), 418–448.

Baranowska-Rataj, A., X. de Luna, and A. Ivarsson (2016). Does the number of siblings affect health in midlife? Evidence from the Swedish prescribed drug register. *Demographic Research 35*(43), 1259–1302.

Barban, N., E. De Cao, S. Oreffice, and C. Quintana-Domeque (2019). Assortative mating on education: A genetic assessment. IZA Discussion Papers 12563, Institute of Labor Economics (IZA).

Barbieri, T., G. Basso, and S. Scicchitano (2020). Italian workers at risk during the covid-19 epidemic. GLO Discussion Paper Series 513, Global Labor Organization (GLO).

Bargain, O. and J. Zeidan (2019). Heterogeneous effects of obesity on mental health: Evidence from Mexico. *Health Economics 28*(4), 447–460.

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

Barro, R. J., J. F. Ursúa, and J. Weng (2020). The coronavirus and the Great Influenza Pandemic: Lessons from the "Spanish flu" for the coronavirus's potential effects on mortality and economic activity. NBER Working Papers 26866, National Bureau of Economic Research, Inc.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy 70*(5), 9–49.

Becker, G. S. and H. G. Lewis (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy 81*(2), 279–288.

Becker, G. S. and Y. Rubinstein (2011). Fear and the Response to Terrorism: An Economic Analysis. CEP Discussion Papers dp1079, Centre for Economic Performance, LSE.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*(6), 1153–1189.

Beland, L.-P., O. Fakorede, and D. Mikola (2020). The Short-Term Effect of COVID-19 on Self-Employed Workers in Canada. GLO Discussion Paper Series 585, Global Labor Organization (GLO).

Bell, B., J. Blundell, and S. Machin (2018). Where is the land of hope and glory? The geography of intergenerational mobility in England and Wales. CEP Discussion Papers 1591, Centre for Economic Performance, LSE.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy 75*(4), 352–365.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender Identity and Relative Income within Households. *The Quarterly Journal of Economics 130*(2), 571–614.

Bhalotra, S. and D. Clarke (2019). The twin instrument: Fertility and human capital investment. *Journal of the European Economic Association 18*(6), 3090–3139.

Bharadwaj, P., M. M. Pai, and A. Suziedelyte (2017). Mental health stigma. *Economic Letters 159*, 57–60.

Bhattacharya, J. and D. Lakdawalla (2006). Does medicare benefit the poor? *Journal of Public Economics 90*(1), 277–292.

Bindler, A., N. Ketel, and R. Hjalmarsson (2020). Costs of victimization. *Handbook of Labor, Human Resources and Population Economics*, 1–31.

Björkegren, E., M. Lindahl, M. Palme, and E. Simeonova (2019). Pre-and post-birth components of intergenerational persistence in health and longevity: Lessons from a large sample of adoptees. *Journal of Human Resources*, online first.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Blanden, J. (2013). Cross-country rankings in intergenerational mobility: A comparison of approaches from economics and sociology. *Journal of Economic Surveys 27*(1), 38–73.

Blau, D. M. (1999). The effect of income on child development. *The Review of Economics and Statistics 81*(2), 261–276.

Blau, D. M. and D. B. Gilleskie (2001). The effect of health on employment transitions of older men. In *Worker Wellbeing in a Changing Labor Market*. Emerald Group Publishing Limited.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature 55*(3), 789–865.

Block, J., A. S. Kritikos, M. Priem, and C. Stiel (2020). Emergency aid for self-employed in the COVID-19 pandemic: A flash in the pan? DIW Discussion Paper 1924, DIW Berlin, German Institute for Economic Research.

Bloom, D. E., E. T. Cafiero, E. Jané-Llopis, S. Abrahams-Gessel, L. R. Bloom, S. Fathima, A. B. Feigl, T. Gaziano, M. Mowafi, A. Pandya, K. Prettner, L. Rosenberg,

B. Seligman, A. Z. Stein, and C. Weinstein (2011). *The Global Economic Burden of Non-communicable Diseases*. Geneva: World Economic Forum.

Blundell, J. and S. Machin (2020). Self-employment in the Covid-19 crisis. CEP Covid-19 Briefings cepcovid-19-003, Centre for Economic Performance, LSE.

Blundell, J. and E. Risa (2019). Income and family background: Are we using the right models? *Available at SSRN 3269576*.

Blundell, R., J. Britton, M. C. Dias, and E. French (2016). The dynamic effects of health on the employment of older workers. Working Papers wp348, University of Michigan, Michigan Retirement Research Center.

Blundell, R., J. Britton, M. C. Dias, and E. French (2021). The impact of health on labor supply near retirement. *Journal of Human Resources*, online first.

Bonacini, L., G. Gallo, and F. Patriarca (2021). Identifying policy challenges of COVID-19 in hardly reliable data and judging the success of lockdown measures. *Journal of Population Economics 34*(1), 275–301.

Bond, T. N. and K. Lang (2018). The black–white education scaled test-score gap in grades K-7. *Journal of Human Resources 53*(4), 891–917.

Bönte, W. and M. Piegeler (2013). Gender gap in latent and nascent entrepreneurship: driven by competitiveness. *Small Business Economics 41*(4), 961–987.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450.

Bratberg, E., J. Davis, B. Mazumder, M. Nybom, D. D. Schnitzlein, and K. Vaage (2017). A comparison of intergenerational mobility curves in Germany, Norway, Sweden, and the US. *The Scandinavian Journal of Economics 119*(1), 72–101.

Bratsberg, B., S. Markussen, O. Raaum, K. Røed, and O. J. Røgeberg (2018). Trends in assortative mating and offspring outcomes. IZA Discussion Papers 11753, Institute of Labor Economics (IZA).

Braun, R. A., K. A. Kopecky, and T. Koreshkova (2017). Old, sick, alone, and poor: A welfare analysis of old-age social insurance programmes. *Review of Economic Studies 84*(2), 580–612.

Briole, S., H. Le Forner, and A. Lepinteur (2020). Children's socio-emotional skills: Is there a quantity–quality trade-off? *Labour Economics*, online first.

Britton, J. and E. French (2020). Health and employment amongst older workers. *Fiscal Studies 41*(1), 221–250.

Brown, R. (2020). The intergenerational impact of terror: Did the 9/11 tragedy impact the initial human capital of the next generation? *Demography 57*, 1459–1481.

Brücker, H., N. Rother, and J. Schupp (2016). *IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse*, Volume 29. Deutsches Institut für Wirtschaftsforschung.

Bundestag (2016). Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Ulla Jelpke, Frank Tempel, Jan van Aken, weiterer Abgeordneter und der Fraktion DIE LINKE. Drucksache 18/7304. Technical report, Deutscher Bundestag - 18. Wahlperiode.

Burda, M. C. (2016). The German labor market miracle, 2003-2015: An assessment. SFB 649 Discussion Paper 2016-005, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk.

Bursztyn, L., G. Egorov, R. Enikolopov, and M. Petrova (2019). Social media and xenophobia: Evidence from Russia. NBER Working Papers 26567, National Bureau of Economic Research, Inc.

Cahuc, P. (2019). Short-time work compensation schemes and employment. *IZA World of Labor*, 1–11.

Cajner, T., L. D. Crane, R. A. Decker, J. Grigsby, A. Hamins-Puertolas, E. Hurst, C. Kurz, and A. Yildirmaz (2020). The U.S. labor market during the beginning of the pandemic recession. NBER Working Papers 27159, National Bureau of Economic Research, Inc.

Caliendo, M., F. M. Fossen, A. S. Kritikos, and M. Wetter (2014). The gender gap in entrepreneurship: Not just a matter of personality. *CESifo Economic Studies 61*(1), 202–238.

Caliendo, M. and S. Künn (2015). Getting back into the labor market: The effects of start-up subsidies for unemployed females. *Journal of Population Economics 28*(4), 1005–1043.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal 17*(2), 372–404.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2362.

Cameron, C. A., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cameron, C. A. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–373.

Carneiro, P., C. Meghir, and M. Parey (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association 11*(s1), 123–160.

Case, A., A. Fertig, and C. Paxson (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics 24*(2), 365–389.

Case, A., D. Lubotsky, and C. Paxson (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review 92*(5), 1308–1334.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2019). Binscatter regressions. *arXiv preprint arXiv:1902.09615*.

Cattaneo, M. D., M. Jansson, and X. Ma (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association 115*(531), 1449–1455.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research 1*(2), 245–276.

Center for the Study of Hate and Extremism (2018). Final U.S. status report hate crime analysis and forecast for 2016/2017. Technical report.

Chalfin, A. (2015). Economic costs of crime. *The encyclopedia of crime and punishment*, 1–12.

Chetty, R., J. N. Friedman, N. Hendren, M. Stepner, and T. O. I. Team (2020). The economic impacts of COVID-19: Evidence from a new public database built using private sector data. NBER Working Papers 27431, National Bureau of Economic Research, Inc.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA 315*(16), 1750–1766.

Chou, S.-Y., J.-t. Liu, M. Grossman, and T. Joyce (2010). Parental education and child health: Evidence from a natural experiment in Taiwan. *American Economic Journal: Applied Economics 2*(1), 33–61.

Clark, A. E., O. Doyle, and E. Stancanelli (2020). The impact of terrorism on individual well-being: Evidence from the boston marathon bombing. *The Economic Journal 130*(631), 2065–2104.

Classen, T. J. and O. Thompson (2016). Genes and the intergenerational transmission of BMI and obesity. *Economics & Human Biology 23*, 121 – 133.

Cobb-Clark, D. A., N. Salamanca, and A. Zhu (2019). Parenting style as an investment in human development. *Journal of Population Economics 32*(4), 1315–1352.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020). Labor markets during the COVID-19 crisis: A preliminary view. NBER Working Papers 27017, National Bureau of Economic Research, Inc.

Coneus, K. and C. K. Spiess (2012). The intergenerational transmission of health in early childhood—evidence from the German Socio-Economic Panel Study. *Economics & Human Biology 10*(1), 89 – 97.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*(3), 79–102.

Corak, M. (2019). The Canadian geography of intergenerational income mobility. *The Economic Journal 130*(631), 2134–2174.

Correia, S., S. Luck, and E. Verner (2020). Pandemics depress the economy, public health interventions do not: Evidence from the 1918 flu. *Available at SSRN 3561560*.

Couch, K. A. and T. A. Dunn (1997). Intergenerational correlations in labor market status: A comparison of the United States and Germany. *The Journal of Human Resources 32*(1), 210–232.

Council of Europe (2016). Violence against migrants. *Report of the Committee on Migration, Refugees and Displaced Persons 14066*.

Cunha, F. (2011). Recent developments in the identification and estimation of production functions of skills. *Fiscal Studies 32*(2), 297–316.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, A., M. A. Shields, and S. W. Price (2007). The child health/family income gradient: Evidence from England. *Journal of Health Economics 26*(2), 213 – 232.

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.

Currie, J. and B. C. Madrian (1999). Health, health insurance and the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 50, pp. 3309–3416. Elsevier.

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

Currie, J. and E. Moretti (2007). Biology as destiny? Short- and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor Economics 25*(2), 231–264.

Currie, J. and M. Stabile (2003). Socioeconomic status and child health: Why is the relationship stronger for older children? *American Economic Review 93*(5), 1813–1823.

Cygan-Rehm, K. (2018). Is additional schooling worthless? Revising the zero returns to compulsory schooling in Germany. CESifo Working Paper 7191, Center for Economic Studies and ifo Institute (CESifo).

Cygan-Rehm, K., D. Kuehnle, and M. Oberfichtner (2017). Bounding the causal effect of unemployment on mental health: Nonparametric evidence from four countries. *Health Economics 26*(12), 1844–1861.

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

Dahl, M. and T. DeLeire (2008). The association between children's earnings and fathers' lifetime earnings: estimates using administrative data. Working Paper 1342-08, Institute for Research on Poverty.

Dahmann, S. C. and D. D. Schnitzlein (2019). No evidence for a protective effect of education on mental health. *Social Science & Medicine 241*, online first.

Dalgaard, C.-J. and H. Strulik (2014). Optimal aging and death: Understanding the Preston Curve. *Journal of the European Economic Association 12*(3), 672–701.

Davidsson, P. and S. R. Gordon (2016). Much ado about nothing? The surprising persistence of nascent entrepreneurs through macroeconomic crisis. *Entrepreneurship Theory and Practice 40*(4), 915–941.

De Nardi, M., S. Pashchenko, and P. Porapakkarm (2017). The lifetime costs of bad health. Working Paper 23963, National Bureau of Economic Research, Inc.

Deaton, A. S. and C. H. Paxson (1998). Aging and inequality in income and health. *American Economic Review 88*(2), 248–253.

Del Boca, D., C. J. Flinn, E. Verriest, and M. J. Wiswall (2019). Actors in the child development process. Working Paper 25596, National Bureau of Economic Research, Inc.

Deole, S. S. (2019). Justice delayed is assimilation denied: Right-wing terror and immigrants' assimilation in germany. *Labour Economics 59*, 69–78.

Deutsche Rentenversicherung (2021). Erwerbsminderungsrenten. `https://www.deutsche-rentenversicherung.de/DRV/DE/Rente/Allgemeine-Informationen/Rentenarten-und-Leistungen/Erwerbsminderungsrente/erwerbsminderungsrente_node.html`. Accessed: 2021-03-17.

Deutsches Institut für Wirtschaftsforschung (2017). IAB-BAMF-SOEP-Befragung Geflüchteter in Deutschland. `https://www.diw.de/de/diw_01.c.538695.de/forschung_beratung/iab_bamf_soep_befragung_gefluechteter_in_deutschland.html`. Accessed: 2017-10-23.

Dingel, J. and B. Neiman (2020). How many jobs can be done at home? *Journal of Public Economics 189*, online first.

Doern, R. (2016). Entrepreneurship and crisis management: The experiences of small businesses during the London 2011 riots. *International Small Business Journal 34*(3), 276–302.

Doern, R., N. Williams, and T. Vorley (2019). Special issue on entrepreneurship and crises: business as usual? An introduction and review of the literature. *Entrepreneurship & Regional Development 31*(5-6), 400–412.

Domnich, A., D. Panatto, R. Gasparini, and D. Amicizia (2012). The "healthy immigrant" effect: Does it exist in Europe today? *Italian Journal of Public Health 9*(3), e7532.

Dustmann, C. (1993). Earnings adjustment of temporary migrants. *Journal of Population Economics 6*(2), 153–168.

Dustmann, C. (1997). Differences in the labor market behavior between temporary and permanent migrant women. *Labour Economics 4*(1), 29–46.

Dustmann, C. (2000). Temporary migration and economic migration. *Swedish Economic Policy Review 7*(2), 213–244.

Dustmann, C. and A. Glitz (2011). Migration and education. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 4, Chapter 4, pp. 327–439. Amsterdam, Netherlands: ElsevierB.V.

Dustmann, C., P. A. Puhani, and U. Schoenberg (2016). The long-term effect of early track choice. *The Economic Journal 127*(603), 1348–1380.

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics 43*, 1–12.

Eika, L., M. Mogstad, and B. Zafar (2019). Educational assortative mating and household income inequality. *Journal of Political Economy 127*(6), 2795–2835.

Elam, A., C. Brush, P. Greene, B. Baumer, M. Dean, and R. Heavlow (2019). Global entrepreneurship monitor 2018/2019 women's entrepreneurship report. *London: Global Entrepreneurship Research Association. Retrieved 25*, 2020.

Entorf, H. and M. Lange (2019). Refugees Welcome? Understanding the Regional Heterogeneity of Anti-Foreigner Hate Crimes in Germany. IZA Discussion Papers 12229, Institute of Labor Economics (IZA).

Fairlie, R. W. (2020). The impact of COVID-19 on small business owners: Evidence from the first three months after widespread social-distancing restrictions. *Journal of Economics & Management Strategy 29*(4), 727–740.

Falk, A., A. Kuhn, and J. Zweimüller (2011). Unemployment and right-wing extremist crime. *The Scandinavian Journal of Economics 113*(2), 260–285.

Federal Ministry for Economic Affairs and Energy (2020). German government announces €50 billion in emergency aid for small businesses. `https://www.bmwi.de/Redaktion/EN/Pressemitteilungen/2020/20200323-50-german-government-announces-50-billion-euros-in-emergency-aid-for-small-businesses.html`, accessed 2020-10-05.

Federal Ministry of Health (2020). Coronavirus SARS-CoV-2: Chronik der bisherigen Maßnahmen. `https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus.html`, accessed 2020-10-21.

Federal Statistical Office of Germany (2019). Empfaenger von Asylbewerberregelleistungen. `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Soziales/Asylbewerberleistungen/Tabellen/4-3-zv-aufenthaltsrechtl-status.html;jsessionid=7336C6BF8A5EA83E1EF7F278E000C0F1.live712`, accessed 2019-10-05.

Fernihough, A. (2017). Less is more? The child quantity-quality trade-off in early 20th century England and Wales. QUCEH Working Paper Series 2017-07, Queen's University, Belfast.

Ferreira, F. H. and V. Peragine (2016). Individual responsibility and equality of opportunity. In *The Oxford handbook of well-being and public policy*.

Fiorillo, D., G. L. Lavadera, and N. Nappo (2020). Structural social capital and mental health: a panel study. *Applied Economics 52*(19), 2079–2095.

Fletcher, J. and K. M. Jajtner (2019). Intergenerational health mobility: Magnitudes and importance of schools and place. Working Paper 26442, National Bureau of Economic Research, Inc.

Forsythe, E., L. B. Kahn, F. Lange, and D. Wiczer (2020). Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims. *Journal of Public Economics 189*, online first.

Foster, G. and L. S. Stratton (2021). Does female breadwinning make partnerships less healthy or less stable? *Journal of Population Economics 34*(1), 63–96.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies 72*(2), 395–427.

French, E. and J. B. Jones (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica 79*(3), 693–732.

Frick, J. R., S. P. Jenkins, D. R. Lilard, O. Lipps, and M. Wooden (2007). The cross-national equivalent file (CNEF) and its member country household panel studies. *Journal of Applied Social Science Studies 127*(4), 627–654.

Friedman, M. (1957). *The Permanent Income Hypothesis*, pp. 20–37. Princeton University Press.

Fritsch, M., A. S. Kritikos, and A. Sorgner (2015). Why did self-employment increase so strongly in Germany? *Entrepreneurship and Regional Development 27*(5-6), 307–333.

Fruehwirth, J. C., S. Iyer, and A. Zhang (2019). Religion and depression in adolescence. *Journal of Political Economy 127*(3), 1178–1209.

Fruewirth, J. C., S. Iyer, and A. Zhang (2019). Religion and depression in adolescence. *Journal of Political Economy 127*(3), 1178–209.

Fujisawa, Y., T. Hamano, and S. Takegawa (2009). Social capital and perceived health in Japan: An ecological and multilevel analysis. *Social Science & Medicine 69*(4), 500 – 505.

Galama, T. J. and H. van Kippersluis (2018). A theory of socio-economic disparities in health over the life cycle. *The Economic Journal*, online first.

Gale, L., W. Heath, and R. Ressler (2002). An economic analysis of hate crime. *Eastern Economic Journal 28*(2), 203–216.

García-Pérez, M. (2016). Converging to American: Healthy immigrant effect in children of immigrants. *American Economic Review 106*(5), 461–66.

Gaudecker, H.-M. v., R. Holler, L. Janys, B. M. Siflinger, and C. Zimpelmann (2020). Labour supply in the early stages of the COVID-19 pandemic: Empirical evidence on hours, home office, and expectations. IZA Discussion Papers 13158, Institute of Labor Economics (IZA).

Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Gelman, A. and G. Imbens (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Georgellis, Y. and H. J. Wall (2005). Gender differences in self-employment. *International Review of Applied Economics 19*(3), 321–342.

Giuntella, O. and F. Mazzonna (2015). Do immigrants improve the health of natives? *Journal of Health Economics 43*, 140–153.

Göbel, J., M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2018). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik 239*(5), 345–360.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik 239*(2), 345 – 360.

Goldin, C., S. P. Kerr, C. Olivetti, and E. Barth (2017). The expanding gender earnings gap: Evidence from the LEHD-2000 census. *American Economic Review: Papers and Proceedings 107*(5), 110–114.

Goldman, D. P. and J. P. Smith (2002). Can patient self-management help explain the SES health gradient? *Proceedings of the National Academy of Sciences 99*(16), 10929–10934.

Gould, E. D. and E. F. Klor (2014). The long-run effect of 9/11: Terrorism, backlash, and the assimilation of Muslim immigrants in the West. *The Economic Journal 126*(597), 2064–2114.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). Marry your like: Assortative mating and income inequality. *American Economic Review 104*(5), 348–53.

Grinde, B. and K. Tambs (2016). Effect of household size on mental problems in children: Results from the Norwegian mother and child cohort study. *BMC Psychology 4*, 31.

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

Günther, L. and K. Marder-Puch (2019). Selbstständigkeit-Methoden und Ergebnisse des Ad-hoc-Moduls zur Arbeitskräfteerhebung 2017. *Wirtschaft und Statistik (WISTA) 1*, 2019.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *The American Economic Review 96*(4), 1308–1320.

Halliday, T., B. Mazumder, and A. Wong (2021). Intergenerational mobility in self-reported health status in the US. *Journal of Public Economics 193*, online first.

Halliday, T. J. and B. Mazumder (2017). An analysis of sibling correlations in health using latent variable models. *Health Economics 26*(12), e108–e125.

Halliday, T. J., B. Mazumder, and A. Wong (2020). The intergenerational transmission of health in the United States: A latent variables analysis. *Health Economics 29*(3), 367–381.

Hausman, C. and D. S. Rapson (2018). Regression discontinuity in time: considerations for empirical applications. *Annual Review of Resource Economics 10*, 533–552.

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *The Journal of Economic Perspectives 15*(4), 57–67.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences 104*(33), 13250–13255.

Helliwell, J. F. and R. D. Putnam (2007). Education and social capital. *Eastern Economic Journal 33*(1), 1–19.

Hendren, N. and B. Sprung-Keyser (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics 135*(3), 1209–1318.

Ho, C. Y. (2016). Better health with more friends: The role of social capital in producing health. *Health Economics 25*(1), 91–100.

Hofmann, S. and A. Mühlenweg (2018). Learning intensity effects in students' mental and physical health – evidence from a large scale natural experiment in Germany. *Economics of Education Review 67*, 216 – 234.

Holmlund, H. (2020). How much does marital sorting contribute to intergenerational socioeconomic persistence? *Journal of Human Resources*, online first.

Hong, S. C. and J. Park (2015). The socioeconomic gradient in the inheritance of longevity: A study of American genealogies. Working papers, Sogang University, Department of Economics.

Hosseini, R., K. A. Kopecky, and K. Zhao (2021). How Important Is Health Inequality for Lifetime Earnings Inequality? FRB Atlanta Working Paper 2021-1, Federal Reserve Bank of Atlanta.

Huebener, M. (2018). The effects of education on health: An intergenerational perspective. IZA Discussion Papers 11795, Institute of Labor Economics (IZA).

Huebener, M. (2019). Life expectancy and parental education. *Social Science & Medicine 232*, 351 – 365.

Hurtado, D., I. Kawachi, and J. Sudarsky (2011). Social capital and self-rated health in Colombia: The good, the bad and the ugly. *Social Science & Medicine 72*, 584–90.

Ifo Institute and forsa (2020). Erste Ergebnisse des Befragungsteils der BMG-Corona-BUND-Studie. `https://www.ifo.de/DocDL/bmg-corona-bund-studie-erste-ergebnisse.pdf`, accessed 2020-10-05.

Islam, M., J. Merlo, I. Kawachi, M. Lindström, and U.-G. Gerdtham (2006). Social capital and health: Does egalitarianism matter? A literature review. *International Journal for Equity in Health 5*(3), 3.

Jasso, G., D. S. Massey, M. R. Rosenzweig, and J. P. Smith (2004). Immigrant health: selectivity and acculturation. In N. B. Anderson, R. A. Bulatao, and B. Cohen (Eds.), *Critical Perspectives on Racial and Ethnic Differences in Health in Late Life*, Chapter 7, pp. 227–266. Washington, DC: National Academy of Sciences.

Johnston, D. W., S. Schurer, and M. A. Shields (2013). Exploring the intergenerational persistence of mental health: Evidence from three generations. *Journal of Health Economics 32*(6), 1077 – 1089.

Juranek, S., J. Paetzold, H. Winner, and F. Zoutman (2020). Labor market effects of COVID-19 in Sweden and its neighbors: Evidence from novel administrative data. Discussion Paper 2020/8, NHH Dept. of Business and Management Science.

Kaiser, H. F. and K. Dickman (1959). Analytic determination of common factors. *American Psychologist 14*(1), 425–441.

Kalenkoski, C. M. and S. W. Pabilonia (2020). Initial impact of the COVID-19 pandemic on the employment and hours of self-employed coupled and single workers by gender and parental status. IZA Discussion Paper 13443, Institute of Labor Economics (IZA).

Karlsson, M., T. Nilsson, and S. Pichler (2014). The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of Health Economics 36*, 1 – 19.

Keane, M., E. Capatina, and S. Maruyama (2018). Health shocks and the evolution of consumption and income over the life-cycle. Discussion Papers 2018-14, School of Economics, The University of New South Wales.

Kemptner, D. (2019). Health-related life cycle risks and public insurance. *Journal of Health Economics 65*, 227–245.

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

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

Kennedy, S., J. T. McDonald, and N. Biddle (2006). The healthy immigrant effect and immigrant selection: Evidence from four countries. Social and Economic Dimensions of an Aging Population Research Papers 164, McMaster University.

Kessler, R. C., G. P. Amminger, S. Aguilar-Gaxiola, J. Alonso, S. Lee, and T. B. Ustun (2007). Age of onset of mental disorders: A review of recent literature. *Current Opinion in Psychiatry 20*(4), 359–364.

Keyes, C. L. M. (2006). Subjective well-being in mental health and human development research worldwide: An introduction. *Social Indicators Research 77*, 1–10.

Khanam, R., H. S. Nghiem, and L. B. Connelly (2009). Child health and the income gradient: Evidence from Australia. *Journal of Health Economics 28*(4), 805–817.

Kim, Y., B. Sikoki, J. Strauss, and F. Witoelar (2015). Intergenerational correlations of health among older adults: Empirical evidence from Indonesia. *The Journal of the Economics of Ageing 6*, 44–56.

Kliem, S., T. Moessle, T. Klatt, S. Fleischer, D. Kudlacek, C. Kroeger, E. Braehler, M. E. Beutel, and W. Joerg (2016). Psychometric evaluation of an arabic version of the PHQ-4 based on a representative survey of syrian refugees. *Psychotherapie, Psychosomatik, medizinische Psychologie 66*(9-10), 385–392.

Koellinger, P., M. Minniti, and C. Schade (2013). Gender differences in entrepreneurial propensity. *Oxford bulletin of Economics and Statistics 75*(2), 213–234.

Kotwal, A. A. (2010). Physical and psychological health of first and second generation Turkish immigrants in Germany. *American Journal of Human Biology 22*(4), 538–545.

Krieger, N., D. R. Williams, and N. E. Moss (1997). Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annual Review of Public Health 18*(1), 341–378.

Kritikos, A. S., D. Graeber, and J. Seebauer (2020). Corona-Pandemie wird zur Krise für Selbständige. *DIW aktuell 47*.

Kroenke, K., R. L. Spitzer, J. B. Wiliams, and B. Loewe (2009). An ultra-brief screening for anxiety and depression: the PHQ-4. *Psychosomatics 50*(6), 613–621.

Krueger, A. B. and J.-S. Pischke (1997). A statistical analysis of crime against foreigners in unified Germany. *The Journal of Human Resources 32*(1), 182–209.

Kuehnle, D. (2014). The causal effect of family income on child health in the UK. *Journal of Health Economics 36*, 137–150.

Kühne, S., M. Kroh, S. Liebig, and S. Zinn (2020). The need for household panel surveys in times of crisis: The case of SOEP-CoV. *Survey Research Methods 14*(2), 195–203.

Lach, S., Y. Ritov, and A. Simhon (2006). Longevity across generations. Discussion Papers 7145, Hebrew University of Jerusalem, Department of Agricultural Economics and Management.

Lampert, T., J. Hoebel, and L. E. Kroll (2019). Social differences in mortality and life expectancy in Germany. Current situation and trends. *Journal of Health Monitoring 4*(1), 3–14.

Landersø, R. and J. J. Heckman (2017). The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US. *The Scandinavian Journal of Economics 119*(1), 178–230.

Layard, R. (2016). The economics of mental health. *IZA World of Labor 321*, 1–10.

Lazear, E. P. (2021). Why are some immigrant groups more successful than others? *Journal of Labor Economics 39*(1), 115–133.

Lee, D. S. and D. Card (2010). Regression discontinuity inference with specification error. *Journal of Econometrics 48*(2), 281–355.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Leoni, T. and M. Falk (2010). Gender and field of study as determinants of self-employment. *Small Business Economics 34*(2), 167–185.

LeSchinsky, A. (1981). *Introduction.* Number 18 in Materialien aus der Bildungsforschung. Berlin.

LeSchinsky, A. and P. M. Roeder (1980). *Didaktik und Unterricht in der Sekundarschule I seit 1950 - Entwicklung der Rahmenbedingungen, Band 1: Entwicklungen seit 1950.* Stuttgart: Klett-Cotta.

Li, H., J. Zhang, and Y. Zhu (2008). The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography 45*(1), 223–243.

Lindeboom, M., A. Llena-Nozal, and B. van der Klaauw (2009). Parental education and child health: Evidence from a schooling reform. *Journal of Health Economics 28*(1), 109 – 131.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *The Review of Economic Studies 72*(1), 189–221.

Lleras-Muney, A. and F. R. Lichtenberg (2005). Are the more educated more likely to use new drugs? *Annales d'Économie et de Statistique*, 671–696.

Loewe, B., I. Wahl, M. Rose, C. Spitzer, H. Glaesmer, K. Wingenfeld, A. Schneider, and E. Braehler (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders 122*(1-2), 86–95.

Long, J. and J. Ferrie (2007). The path to convergence: Intergenerational occupational mobility in britain and the us in three eras. *The Economic Journal 117*(519), C61–C71.

Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review 103*(4), 1109–37.

Luft, S. (2016). *Die Flüchtlingskrise: Ursachen, Konflikte, Folgen.* CH Beck.

Lundborg, P., A. Nilsson, and D.-O. Rooth (2014). Parental education and offspring outcomes: Evidence from the Swedish compulsory school reform. *American Economic Journal: Applied Economics 6*(1), 253–78.

Lynch, J. W., G. A. Kaplan, and S. J. Shema (1997). Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. *New England Journal of Medicine 337*(26), 1889–1895.

Marcus, J. (2013). The effect of unemployment on the mental health of spouses - Evidence from plant closures in Germany. *Journal of Health Economics 32*(3), 546 – 558.

Margaryan, S., A. Paul, and T. Siedler (2019). Does education affect attitudes towards immigration? Evidence from Germany. *Journal of Human Resources*, online first.

Markussen, S. and K. Røed (2020). Economic mobility under pressure. *Journal of the European Economic Association 18*(4), 1844–1885.

Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *The Review of Economics and Statistics 87*(2), 235–255.

McCrary, B. J. and H. Royer (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review 101*(2), 158–195.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698 – 714.

Meara, K., F. Pastore, and A. Webster (2020). The gender pay gap in the USA: a matching study. *Journal of Population Economics 33*(1), 271–305.

Metcalfe, R., N. Powdthavee, and P. Dolan (2011). Destruction and distress: Using a quasi-experiment to show the effects of the September 11 attacks on mental well-being in the United Kingdom. *The Economic Journal 121*, F81–F103.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology 50*(5), 517 – 528.

Milani, F. (2021). COVID-19 outbreak, social response, and early economic effects: a global VAR analysis of cross-country interdependencies. *Journal of Population Economics 34*(1), 223–252.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy 66*(4), 281–281.

Modalsli, J. (2017). Intergenerational mobility in Norway, 1865–2011. *The Scandinavian Journal of Economics 119*(1), 34–71.

Mogstad, M. and M. Wiswall (2016). Testing the quantity–quality model of fertility: Estimation using unrestricted family size models. *Quantitative Economics 7*(1), 157–192.

Müller, K. and C. Schwarz (2020). From hashtag to hate crime: Twitter and anti-minority sentiment. *Available at SSRN 3149103*.

Müller, K. and C. Schwarz (2020). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, online first.

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

Niemeyer, H., A. Bieda, J. Michalak, S. Schneider, and J. Margraf (2019). Education and mental health: Do psychosocial resources matter? *SSM - Population Health 7*, online first.

Nuebling, M., H. H. Andersen, A. Muehlbacher, J. Schupp, and G. Wagner (2007). Computation of standard values for physical and mental health scale scores using the SOEP version of SF12v2. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift fuer Wirtschafts- und Sozialwissenschaften 127*(1), 171–182.

Nybom, M. and J. Stuhler (2016). Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources 51*(1), 239–268.

Nybom, M. and J. Stuhler (2017). Biases in standard measures of intergenerational income dependence. *Journal of Human Resources 52*(3), 800–825.

OECD (2017). Entrepreneurship at a glance. `https://www.oecd-ilibrary.org/docserver/entrepreneur_aag-2017-en.pdf?expires=1610536834&id=id&accname=guest&checksum=D90265D7A9E5AB441BB39E57CA65F57A`, accessed 2021-01-13.

Organization for Economic Co-operation and Development (2017). Finding their way - labour market integration of refugees in Germany. Technical report, OECD.

O'Rourke, H. and D. MacKinnon (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of Studies on Alcohol and Drugs 79*, 171–181.

Palloni, A. and E. Arias (2004). Paradox lost: Explaining the hispanic adult mortality advantage. *Demography 41*(3), 385–415.

Papageorge, N. W., G. C. Pauley, M. Cohen, T. E. Wilson, B. H. Hamilton, and R. A. Pollak (2019). Health, human capital and domestic violence. *Journal of Human Resources*, online first.

Parker, S. C. (2018). *The economics of entrepreneurship* (2nd ed.). Cambridge: Cambridge University Press.

Pascual, M. and D. Cantarero (2009). Intergenerational health mobility: An empirical approach based on the ECHP. *Applied Economics 41*(4), 451–458.

Patrick, C., H. Stephens, and A. Weinstein (2016). Where are all the self-employed women? Push and pull factors influencing female labor market decisions. *Small Business Economics 46*(3), 365–390.

Perry, M. (1996). The relationship between social class and mental disorder. *The Journal of Primary Prevention 17*(1), 17–30.

Persson, P. and M. Rossin-Slater (2018). Family ruptures, stress, and the mental health of the next generation. *American Economic Review 108*(4-5), 1214–52.

Petrou, S. and E. Kupek (2008). Social capital and its relationship with measures of health status: Evidence from the health survey for England 2003. *Health Economics 17*(1), 127–143.

Petrovan, S., D. Aldridge, H. Bartlett, A. Bladon, H. Booth, S. Broad, D. Broom, N. Burgess, A. Cunningham, M. Ferri, A. Hinsley, A. Hughes, K. Jones, M. Kelly, G. Mayes, C. Ugwu, N. Uddin, D. Verissimo, T. White, and W. Sutherland (2020). Post COVID-19: a solution scan of options for preventing future zoonotic epidemics.

Petzold, H.-J. (1981). *Schulzeitverlängerung oder Bildungschance? Die Funktion des 9. und 10. Bildungsjahres.* Bamberg: päd.extra buchverlag.

Pijoan-Mas, J. and J.-V. Ríos-Rull (2014). Heterogeneity in expected longevities. *Demography 51*(6), 2075–2102.

Piopiunik, M. (2014). Intergenerational transmission of education and mediating channels: Evidence from a compulsory schooling reform in Germany. *The Scandinavian Journal of Economics 116*(3), 878–907.

Pischke, J.-S. and T. von Wachter (2008). Zero returns to compulsory schooling in Germany: Evidence and interpretation. *The Review of Economics and Statistics 90*(3), 592–598.

Qiu, Y., X. Chen, and W. Shi (2020). Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics 33*(4), 1127–1172.

Ravesteijn, B., H. v. Kippersluis, and E. v. Doorslaer (2018). The wear and tear on health: What is the role of occupation? *Health Economics 27*(2), e69–e86.

Rawlings, S. B. (2015). Parental education and child health: Evidence from an education reform in China. CINCH Working Paper Series 1511, Universitaet Duisburg-Essen, Competent in Competition and Health.

Razum, O., H. Zeeb, H. S. AkgÃ¼n, and S. Yilmaz (1998). Low overall mortality of Turkish residents in Germany persists and extends into a second generation: Merely a healthy migrant effect? *Tropical Medicine & International Health 3*(4), 297–303.

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

Ronda, V. (2016). The effect of maternal psychological distress on children's cognitive development. mimeo.

Roodman, D., J. G. MacKinnon, M. Ørregaard Nielsen, and M. D. Webb (2019). Fast and wild: Bootstrap inference in Stata using boottest. *Stata Journal 19*(1), 4–60.

Rosenzweig, M. and K. I. Wolpin (1980). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica 48*(1), 227–40.

Ruby, J. G., K. M. Wright, K. A. Rand, A. Kermany, K. Noto, D. Curtis, N. Varner, D. Garrigan, D. Slinkov, I. Dorfman, J. M. Granka, J. Byrnes, N. Myres, and C. Ball (2018). Estimates of the heritability of human longevity are substantially inflated due to assortative mating. *Genetics 210*(3), 1109–1124.

Salyers, M. P., H. B. Bosworth, J. W. Swanson, J. Lamb-Pagone, and F. C. Osher (2000). Reliability and validity of the SF-12 health survey among people with severe mental illness. *Medical Care 38*(11), 1141–1150.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement 34*(4, Pt. 2), 1–97.

Schilbach, F., H. Schofield, and S. Mullainathan (2016). The psychological lives of the poor. *American Economic Review 106*(5), 435–40.

Schnitzlein, D. D. (2016). A new look at intergenerational mobility in Germany compared to the U.S. *Review of Income and Wealth 62*(4), 650–667.

Schwarze, J., H. H. Andersen, and S. Anger (2000). Self-rated health and changes in self-rated health as predictors of mortality: First evidence from German panel data. DIW Discussion Papers 203, DIW Berlin, German Institute for Economic Research.

Sen, A. (2002). Why health equity? *Health Economics 11*(8), 659–666.

Siemer, M. (2014). Firm entry and employment dynamics in the great recession. Finance and Economics Discussion Series 2014-56, Board of Governors of the Federal Reserve System (U.S.).

Silles, M. A. (2015). The intergenerational effect of parental education on child health: Evidence from the UK. *Education Economics 23*(4), 455–469.

Simon, G., M. Korff, K. Saunders, D. Miglioretti, P. Crane, G. Belle, and R. Kessler (2006). Association between obesity and psychiatric disorders in the US adult population. *Archives of general psychiatry 63*, 824–30.

Social Security Disability Insurance (2021). Substantial gainful activity. `https://www.ssa.gov/oact/cola/sga.html`. Accessed: 2021-03-17.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review 82*(3), 393–408.

Solon, G. (1999). Intergenerational mobility in the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 29, pp. 1761–1800. Elsevier.

Sorgner, A., M. Fritsch, and A. S. Kritikos (2017). Do entrepreneurs really earn less? *Small Business Economics 49*(2), 251–272.

Staiger, B. Y. D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Steinhardt, M. F. (2018). The impact of xenophobic violence on the integration of immigrants. IZA Discussion Papers 11781, Institute of Labor Economics (IZA).

Steinmetz, H., V. Batzdorfer, and M. Bosnjak (2020). The ZPID lockdown measures dataset for germany. Technical report, PsychArchives.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In J. H. Stock and D. W. K. Andrews (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, pp. 80—-108. Cambridge: Cambridge University Press.

Stoica, O., A. Roman, and V. D. Rusu (2020). The nexus between entrepreneurship and economic growth: A comparative analysis on groups of countries. *Sustainability 12*(3), 1–1.

Tanner, J., H. Goldstein, and R. Whitehouse (1970). Standards for children's height at ages 2-9 years allowing for height of parents. *Archives of disease in childhood 45*(244), 755–762.

Tent, J. F. (1982). Mission on the Rhine: American educational policy in postwar Germany, 1945–1949. *History of Education Quarterly 22*(3), 255–276.

Ther, P. (2019). *The Outsiders: Refugees in Europe Since 1492*. Princeton University Press.

Thomas, D., J. Strauss, and M.-H. Henriques (1991). How does mother's education affect child height? *The Journal of Human Resources 26*(2), 183–211.

Thompson, O. (2017). Gene-environment interaction in the intergenerational transmission of asthma. *Health Economics 26*(11), 1337–1352.

Timothy, H. (2011). Health inequality over the life-cycle. *The B.E. Journal of Economic Analysis & Policy 11*(3), 1–21.

TNS Infratest Sozialforschung (2014). *Erhebungsinstrumente 2014 (Welle 31) des Sozio-oekonomischen Panels: Personenfragebogen, Altstichproben.* Berlin: DIW/SOEP.

Ullmann, S. H., N. Goldman, and D. S. Massey (2011). Healthier before they migrate, less healthy when they return? The health of returned migrants in Mexico. *Social Science & Medicine 73*(3), 421 – 428.

UNHCR (2009). Global trends 2008 - refugees, asylum-seekers, returnees, internally displaced and stateless persons. Technical report, UNHCR.

UNHCR (2019). Global trends 2018 - forced displacement in 2018. Technical report, UNHCR.

van Doorslaer, E. and U.-G. Gerdtham (2003). Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. *Social Science & Medicine 57*(9), 1621 – 1629.

Velde, F. R. (2020). What happened to the US economy during the 1918 Influenza pandemic? A view through high-frequency data. Working Paper WP 2020-11, Federal Reserve Bank of Chicago.

Vilagut, G., C. G. Forero, A. Pinto-Meza, J. M. Haro, R. D. Graaf, R. Bruffaerts, V. Kovess, G. D. Girolamo, H. Matschinger, J. Alonso, and E. Investigators (2013). The mental component of the Short-Form 12 health survey (SF-12) as a measure of depressive disorders in the general population: Results with three alternative scoring methods. *Value in Health 16*(4), 564–573.

Wagner, S., D. Boertien, and M. Goertz (2020). The wealth of parents: Trends over time in assortative mating based on parental wealth. CEBI working paper series 19-03, The Center for Economic Behavior and Inequality, University of Copenhagen.

Wang, H., X. Yang, T. Yang, R. Cottrell, L. Yu, X. Feng, and S. Jiang (2015). Socioeconomic inequalities and mental stress in individual and regional level: A twenty one cities study in China. *International Journal for Equity in Health 14*(1).

Wegener, B. (1988). *Kritik des Prestiges*. Opladen: Westdeutscher Verlag.

Weiss, Y. (1993). The formation and dissolution of families: Why marry? Who marries whom? And what happens upon divorce. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1 of *Handbook of Population and Family Economics*, Chapter 3, pp. 81–123. Elsevier.

Willage, B. (2018). The effect of weight on mental health: New evidence using genetic IVs. *Journal of Health Economics 57*, 113 – 130.

Wong, P. K., Y. P. Ho, and E. Autio (2005). Entrepreneurship, innovation and economic growth: Evidence from GEM data. *Small Business Economics 24*(3), 335–350.

World Health Organization (2020). Obesity. `https://www.who.int/topics/obesity/en/`, accessed 2020-05-25.

Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *The American Economic Review 82*(3), 409–429.

zur Nieden, F. and B. Sommer (2016). Life expectancy in Germany based on the 2011 census: Was the healthy migrant effect merely an artefact? *Comparative Population Studies 41*(2), 145–174.