SCHOOL OF
ECONOMICS AND
MANAGEMENT

Master's Programme in Economics

# The Role of Education in Health Behaviors

An IV Study Using the SHARE Data

Elvira Wiktorsson

NEKP01
Second Year Master's Thesis (15 credits ECTS)
June 2022
Supervisor: Therese Nilsson

# Abstract

The association between education and health is well documented: more educated individuals live longer and have healthier lives. The incidence of health disparities by education was further accentuated during the Covid-19 pandemic, where the rates of Covid-19 cases and fatalities were higher in communities with less-educated populations. Differences in education, and related differences in health behaviors, might be a mechanism through which these health disparities run. This paper examines the causal effect of education on health behaviors among people over 50 years in Europe using data from the *Survey of Health, Ageing and Retirement in Europe*. I use an instrumental variable approach, exploiting exogenous variation in compulsory education stemming from compulsory schooling reforms in European countries. First, I study the causal effect of education on the probability of smoking, alcohol consumption, engagement in physical activity and having a healthy diet. Second, I examine the effect of education on three Covid-19 related health behaviors: washing hands, using of facemask and preferences to get vaccinated against Covid-19. The results indicate that one additional year of education decreases alcohol consumption and the likelihood of having ever smoked. These negative effects of education seem to be driven by women. In contrast, one more year of education has a negative effect on the consumption of fruits and vegetables and no significant effect on physical activity and Covid-19 related health behaviors.



**Key words:** Education, Health behaviors, Covid-19, Instrumental variable approach, Compulsory schooling reforms, Health economics, Education economics.

# Contents

**1 Introduction** **5**

**2 Previous literature** **7**
2.1 The role of health behaviors in the education-health gradient . . . . . . . . 7
2.2 The effect of education on health behaviors . . . . . . . . . . . . . . . . . . . 7
2.2.1 Smoking behavior . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.2.2 Physical activity . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
2.2.3 Diet and alcohol consumption . . . . . . . . . . . . . . . . . . . . . 9
2.2.4 Education and Covid-19 . . . . . . . . . . . . . . . . . . . . . . . . 10
2.2.5 Heterogeneous effects across gender . . . . . . . . . . . . . . . . . . 10

**3 Theoretical framework** **11**
3.1 Education and health . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
3.2 The effect of compulsory school reforms on schooling . . . . . . . . . . . . . 12
3.3 The effect of compulsory school reforms on health behavior . . . . . . . . . 13

**4 Data** **14**
4.1 SHARE data . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
4.1.1 Years of education . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
4.1.2 Main health behaviors . . . . . . . . . . . . . . . . . . . . . . . . . 16
4.1.3 Covid-19 related health behaviors . . . . . . . . . . . . . . . . . . . 16
4.2 Compulsory schooling reforms . . . . . . . . . . . . . . . . . . . . . . . . . 16
4.3 Descriptive statistics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

**5 Empirical approach** **18**
5.1 Empirical specification . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
5.2 The linear probability model . . . . . . . . . . . . . . . . . . . . . . . . . . 20
5.3 The IV-approach and the LATE estimate . . . . . . . . . . . . . . . . . . . 21
5.3.1 Assumptions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
5.3.2 The LATE estimate . . . . . . . . . . . . . . . . . . . . . . . . . . . 22

**6 Selection of reforms** **23**

**7 Results** **25**
7.1 First stage results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
7.1.1 Visual presentation of the first stage . . . . . . . . . . . . . . . . . 25
7.1.2 Event study results . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
7.2 Second stage results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
7.2.1 Main health behaviors . . . . . . . . . . . . . . . . . . . . . . . . . 29

7.2.2 Covid-19 related health behaviors . . . . . . . . . . . . . . . . . . . . 30
7.3 Heterogeneous effects . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
7.3.1 Main health behaviors . . . . . . . . . . . . . . . . . . . . . . . . 34
7.3.2 Covid-19 related health behaviors . . . . . . . . . . . . . . . . . . . . 35
7.4 Robustness . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38

**8 Discussion** **41**
8.1 Relation to previous studies and the theoretical framework . . . . . . . . . . 41
8.2 Possible explanations for the variation in the results . . . . . . . . . . . . . . 42
8.3 Heterogeneous effects . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43

**9 Conclusion** **44**

**10 References** **46**

**A Appendix** **51**

# List of Tables


1 Descriptive statistics . . . . . . . . . . 18
2 Selected compulsory schooling reforms . . . . . . . . . . 25
3 Main health behaviors . . . . . . . . . . 32
4 Covid-19 related health behaviors . . . . . . . . . . 33
5 Heterogenous effects - Main health behaviors . . . . . . . . . . 36
6 Heterogenous effects - Covid-19 related health behaviors . . . . . . . . . . 37
7 Robustness test - Main health behaviors . . . . . . . . . . 39
8 Robustness test - Covid-19 related health behaviors . . . . . . . . . . 40
A.1 Compulsory schooling reforms . . . . . . . . . . 51
A.2 Selection of reforms . . . . . . . . . . 53
A.3 Descriptive statistics by gender - Main sample . . . . . . . . . . 54
A.4 Descriptive statistics by gender - Covid-19 sample . . . . . . . . . . 54


# List of Figures


1 Main sample - Increasing compulsory education . . . . . . . . . . 26
2 Main sample - Decreasing compulsory education . . . . . . . . . . 26
3 Covid-19 sample - Increasing compulsory education . . . . . . . . . . 28
4 Covid-19 sample - Decreasing compulsory education . . . . . . . . . . 28
5 Main sample - Event study . . . . . . . . . . 29
6 Covid-19 sample - Event study . . . . . . . . . . 29
A.1 First stage by reform . . . . . . . . . . 52

# 1 Introduction

Education is a main determinant of health. In the United States, individuals at the age of 25 that have more education than a college degree, can expect to live up to seven years longer than those who do not hold a college degree. These health disparities by education are observed in many countries and seem to be growing over time (Meara et al., 2008; Hummer and Hernandez, 2013).

A major reason for these health disparities is differences in health behavior (Cutler and Lleras-Muney, 2010). More educated individuals smoke less, drink less alcohol, have healthier diets, and exercise more (Galama et al. 2018). Given that smoking, obesity and alcohol consumption are leading causes of preventable death and disease, it is essential to understand the role of education for unhealthy behaviors (Mokdad et al., 2004). If the effect of education on health behaviors is large enough, then education policies might be crucial measures to limit health disparities.

The well-known health-education gradient has been further accentuated in the Covid-19 pandemic. Communities with less-educated populations experienced substantially higher rates of Covid-19 cases and fatalities during the pandemic (Hawkins et al., 2020). This naturally raises the question of whether these disparities are due to differences in health behavior related to education? Disregarding behaviors such as using facemask, washing hands and getting vaccinated against Covid-19, was risky during the pandemic, especially for individuals at a higher age facing a higher risk of serious complications of a Covid-19 infection. Investigating the causal effect of education on these health behaviors is important to understand if education policies can limit health disparities in *at-risk* situations, such as a pandemic.

While the association between education and health behaviors is well documented, the results of the causal effect of education on health behaviors are mixed. In other words, studies that try to take endogeneity into account find results that vary over health behaviors, country and time period (Galama et al., 2018). In an attempt to reduce uncertainty and improve knowledge in this area of research, the aim of this thesis is to examine the causal effect of education on risky health behaviors among people over 50 years in Europe. I investigate whether there is a causal effect of education on the probability of smoking, alcohol consumption, engagement in physical activity and having a healthy diet. In addition, I study the effect of education on risky health behaviors in the setting of the Covid-19 pandemic, focusing on three health behaviors: washing hands, using of facemask and wanting to get vaccinated against Covid-19. Finally, I examine the potential heterogeneous effects of education on health behaviors by gender.

To identify the causal effect of education on health behavior, I use an instrumental variable (IV) approach. In line with Brunello et al. (2013a) and Hofmarcher (2021), I exploit variation in compulsory education stemming from compulsory schooling reforms in Europe during the 20th century. I use individual data on health, health behaviors and living conditions for individuals over 50 years in Europe from the *Survey of Health, Ageing and Retirement in Europe* (SHARE) combined with Hofmarcher's (2021) database on compulsory schooling reforms in Europe.

My study contributes to the existing literature in two main ways. First, by examining the causal effect of education on a set of health behaviors using a multi-country setup. I apply the well-established method of taking advantage of variation in compulsory education due to compulsory schooling reforms to a multi-country setting instead of investigating the effect within a single country. This adds to the existing literature by presenting more general results of the effect of education on health behaviors, that are not specific to a certain country, time period or population. Second, this thesis contributes to the literature covering the importance of education in the Covid-19 pandemic. Previous literature examines the association between education and Covid-19 health outcomes, but not health behaviors. This paper is thus an important contribution since it tries to understand the mechanism behind this association, by examining the causal effect of education on Covid-19 related health behaviors.

The results suggest that one additional year of education decreases alcohol consumption and the likelihood of having ever smoked daily, while it has a small but negative effect on the consumption of fruits and vegetables. No significant effect of education is found on physical activity or health behaviors during the Covid-19 pandemic. The results differ in magnitude across gender and the negative effect of education on smoking and alcohol consumption seems to be driven by women. The results are robust to changes in the sample bandwidth. I conclude that education policies can be used as a measure to limit disparities in some health behaviors, but not in others. While education seems to lead to individuals avoiding unhealthy consumption, it does not seem to encourage the adoption of healthy behaviors. Education does therefore not seem to affect the adoption of protective behaviors in an *at-risk* situation such as the Covid-19 pandemic.

# 2 Previous literature

## 2.1 The role of health behaviors in the education-health gradient

The association between education and health is well documented in the literature. Individuals with more education live longer and are healthier during their lifetime, reporting fewer health conditions than less educated individuals (Galama et al., 2018; Cutler and Lleras-Muney, 2014). Empirical studies also find strong associations between education and health behaviors. Among more educated individuals, smoking prevalence is lower, consumption of alcohol is lower, exercise is more frequent and the diet is more healthy (Droomers et al., 1999; Cutler and Lleras-Muney, 2010; Cutler and Glaeser, 2005).

Less is known about whether differences in health behaviors by education can explain the relationship between education and health. There is however some evidence suggesting that health behaviors constitute an important factor in the education-health gradient. Brunello et al. (2016) find that education has a positive effect on health and that smoking, drinking, exercising and BMI play a role in explaining the relationship between education and health. They also find that health behaviors can explain a larger share of the effect of education on health when they consider the history of health behaviors instead of only considering recent health behaviors. Health behaviors can explain 23 to 45 percent of the effect of education on health when considering the history of health behaviors, while it only explains 17 to 31 percent when considering recent health behaviors (Brunello et al., 2016).

## 2.2 The effect of education on health behaviors

Recent empirical work has tried to address whether the relation between education and health behaviors is causal, but the results are mixed (Cutler and Lleras-Muney, 2014; Galama et al., 2018). The effect of education on specific health behaviors has been investigated with several empirical approaches. A few studies use randomized control trials (RCT) to investigate a potential causal effect of education on health behaviors (see for example Heckman et al., 2013 and Conti et al., 2016). But since there only exist a few RCTs for childhood education, the large majority of studies try to encounter the endogeneity problem by estimating the effect of education on health by performing twin studies or with quasi-experimental approaches. A common approach in the empirical literature has been to use an IV approach or a regression discontinuity (RD) design by taking advantage of exogenous variation in compulsory schooling due to compulsory schooling reforms, similar to what is done in this paper (Galama et al., 2018; Cutler and Lleras-Muney, 2014). In the following sections, I present results from the existing literature covering the health behaviors that will be examined in this thesis.

### 2.2.1 Smoking behavior

The results from the empirical literature on the effect of education on smoking are mixed. The results from the RCT studies by Heckman et al. (2013) and Conti et al. (2016) suggest that males that were assigned to preschool education in the Perry Preschool Program in the 1960s in the U.S. were less likely to smoke both at the age of 27 and the age of 40. These males also had a lower lifetime prevalence of smoking at the age of 40. No such effects were found among women assigned to preschool education in the same program.

In contrast to this, evidence from twin studies suggests that education does not affect smoking behavior. Neither Lundborg (2013), using data on twins in the U.S., nor Amin et al. (2013), using UK twin data, find a significant effect of education on smoking. In line with these results, the majority of studies using a quasi-experimental approach find no effect of education on smoking behavior. Using an RD design, Clark and Royer (2013) take advantage of compulsory schooling reforms in the UK and find no effect of education on smoking prevalence or for having ever smoked. Exploiting exogenous variation in compulsory schooling in an IV approach, Braakman (2011) finds no significant effect of education on smoking in England, Kemptner et al. (2011) find no effect on currently smoking or on the likelihood of having ever smoked for both men and women in West Germany and Park and Kang (2008) find no evidence of an effect of education on smoking prevalence among Korean men.

There is however some quasi-experimental evidence of a causal effect of education on smoking. Etilé and Jones (2011) study the effect of an education expansion in France using a Difference-in-Difference (DiD) approach. Their results suggest that one year of extra schooling decreases smoking prevalence by 6 to 8 percent. They also find that more education increases the chance of quitting smoking at any given age and decreases the hazard of starting smoking.

### 2.2.2 Physical activity

The literature on the effect of education on physical activity is limited, but the majority of the existing studies find evidence that education increases physical activity. In his U.S. twin study, Lundborg (2013) find a positive effect of education on physical activity. The rest of the evidence is from studies using a quasi-experimental approach. Atella and Kopinska (2014) use an IV approach and find a positive effect of education on physical activity among women in Italy. Li and Powdthavee (2015), exploit changes in compulsory schooling laws in Australia and show that more education increases the likelihood of regular exercise. Park and Kang (2008) find that one more year of schooling on average increases the probability of regular exercise by 7 to 11 percentage points among Korean men. There is however some conflicting evidence, Clark and Royer (2013) find no significant effect of education on exercise in the U.K. in their study using an RD approach.

### 2.2.3 Diet and alcohol consumption

There is some evidence from quasi-experimental studies that more education leads to a more healthy diet. Atella and Kopinska (2014) find evidence of education promoting a more balanced diet by decreasing caloric intake. Li and Powdthavee (2015) also find a positive effect of education on people's diets. In contrast, Clark and Royer (2013) find no effect of education on a healthy diet and Braakman (2011) finds no effect of education on the consumption of fruits and vegetables.

The evidence for the effect of education on alcohol consumption is also mixed. The RCT study by Conti et al. (2016) find that women treated by the Perry Preschool Program drink less frequently at the age of 27, but the effect is no longer significant at the age of 40. The evidence from quasi-experimental studies however suggests that there is no effect of education on alcohol consumption. Clark and Royer (2013) find no significant results for the effect of education on drinking, Braakman (2011) finds no evidence of education affecting excessive drinking and Park and Kang (2008) find no causal effect of education on drinking moderately.

The empirical literature examining the effect of education on BMI, overweight or obesity is broader than the literature examining drinking behavior or diet. If BMI, overweight and obesity are negatively affected by education, drinking behavior and diet are possible mechanisms through which this effect runs. There is some evidence of a protective effect of education on obesity from RCT studies. Conti et al. (2016) show that individuals that were offered early childhood education in the Abecedarian Program in the U.S. in the 1970s, were less likely to be obese, but this effect is not significant and should be considered with caution since the program also included health services and nutritional support. It is therefore not sure whether the effect stemmed only from increased education (Galama et al. 2018). While the U.S. twin study by Lundborg (2013) find no effect of education on BMI or overweight, the quasi-experimental evidence is in line with the results in Conti. et al. (2016). Brunello et al. (2013a) take advantage of compulsory schooling reforms in Europe to estimate the effect of education on obesity, measured as BMI. The authors find that education has a protective effect on BMI, overweight and obesity for women, but not for men. Similarly, Atella and Kopinska (2014) find a negative effect of education on BMI for women, but not for men.

### 2.2.4 Education and Covid-19

The Covid-19 pandemic has shed additional light on the relationship between education and health. At the onset of the pandemic, researchers quickly identified age and comorbidity as risk factors for mortality and serious complications due to a Covid-19 infection, but socioeconomic factors also seem to play a role in the risk of serious illness (Hawkins et al., 2020). Hawkins et al. (2020) study the relationship between socioeconomic factors and disparities in the prevalence and mortality of Covid-19 at the community level in the U.S.. They find that lower education, measured as the percentage of adults without a high school degree, is one of two socioeconomic determinants of health that are strongest positively associated with both Covid-19 cases and fatalities. The other socioeconomic determinant strongly positively associated with Covid-19 cases and fatalities is the percentage of black residents.

Hawkins et al. (2020) investigate the relationship between education and Covid-19 related health outcomes, but it does not examine the effect of health behaviors. Differences in health behaviors due to educational differences could potentially explain the strong relationship between education and Covid-19 cases and fatalities that has been documented.

### 2.2.5 Heterogeneous effects across gender

Galama et al. (2018) report that many studies that estimate the effect of education on different health behaviors find that the results differ across gender (see for example Brunello et al, 2013a, Brunello et al., 2016, and Atella and Kopinska, 2014) and that the effects, in general, are weaker for women than for men. A potential reason for this weaker effect might be pregnancy, which correlates with changes in health-seeking behavior and laborforce participation. For example, women with small children make more use of the healthcare system and are strongly advised not to smoke during pregnancy. These facts are true for women with all levels of education and might attenuate the effect of education on health behavior (Galama et al., 2018). The negative impact of having children on women's labor supply and hourly earnings (Lundborg et al., 2017), might be another explanatory factor for the weaker effect. This is since it weakens the relationship between education and lifetime earnings for women and thus also between education and health behaviors (Galama et al., 2018).

Some health behaviors such as smoking and consumption of alcohol were regarded as taboo for women at the beginning of the 20th century, but as women entered the labor force, these health behaviors became more accepted and even a symbol of independence. Consequently, for women in more recent decades, a potential negative effect of education on smoking or alcohol consumption might be attenuated by a normative incentive for these unhealthy behaviors among successful women in the labor market (Galama et al., 2018; Amos and Haglund, 2000).

# 3 Theoretical framework

The theoretical framework presented below outlines the mechanisms behind the association between education and health, in which health behaviors play an important role. This framework highlights efficiency as the main mechanism behind the fact that more educated people tend to engage in more healthy behavior and thus have better health. This efficiency mechanism can either be due to productive efficiency or due to allocative efficiency. According to the productive efficiency theory, education raises the productivity in the individual's production of health capital. The allocative efficiency theory is instead based on the idea that more educated individuals are more effective in allocating health inputs in the production of health (Grossman, 2006).

## 3.1 Education and health

Grossman (1972) developed a theoretical framework in which individuals produce health as capital and where human capital is taken into account in this production. His work constitutes the base of the productive efficiency theory. In the Grossman model, consumers both demand and produce health. They demand health for two reasons. First, health is seen as a consumption commodity since good health increases utility. Second, health is seen as an investment commodity since it determines the time the individual can use for market or nonmarket activities. Health capital is seen as a part of an individual's human capital. An individual is born with an initial amount of health capital. The health capital then decreases with age and can be affected by investment in health. Investment in health is done by adding health inputs to the production of health. Such health inputs are for example use of medical care, a healthy diet, physical activity, smoking and alcohol consumption. Efficiency, the amount of health capital obtained from a certain amount of health inputs, varies across consumers and affects their health production function. The consumer's efficiency is decided by for example the number of years in education (Grossman, 2000). In this model, more educated individuals consequently have better health due to a more efficient health production function.

Galama et al. (2018) develop the Grossman model by incorporating the allocative efficiency theory into this model of productive efficiency and making the schooling decision endogenous. In the Grossman model, longevity is fixed and the years of education are exogenous (Galama et al., 2018). Galama et al. (2018) extend the model in Grossman (1972) by treating health, skills, health behavior, schooling and longevity as endogenous variables. By making health behaviors endogenous in the model, they incorporate the allocative efficiency theory into Grossman's model. Making schooling endogenous also allows for taking into account the effect of laws and institutions on education. This is useful since it makes it possible to use the model to predict the effects of compulsory schooling reforms on the total years of education.

In the model by Galama et al. (2018), human capital is divided into two components: health and skills. The latter includes both cognitive and non-cognitive skills. Individuals can produce health by consumption of inputs such as food and medical care, while they can increase skills by making investments such as schooling. These consumption and investment choices made by the individual will affect skills and health directly, which in turn will affect labor market outcomes and consumption choices and thereby also health and mortality. Individuals strive to maximize their lifetime utility and they get utility from consumption and health. In contrast, they get disutility from spending time in school. The stock of health depreciates, but the depreciation rate depends on the individual's stock of health, age, consumption and endowments. Consumption of goods and services can be both healthy and unhealthy. Healthy consumption is for example consumption of fruits and vegetables and unhealthy consumption is for example smoking. Both types of consumption create utility, but while healthy consumption negatively affects the health depreciation rate, unhealthy consumption increases the health depreciation rate.

## 3.2 The effect of compulsory school reforms on schooling

In the model by Galama et al. (2018), the optimal choice of schooling and consumption depends on the relation between the benefits of staying in school and the benefits of entering the labor market. The individual will enter the labor market at the age when the net benefits of working exceed the net benefits of continued education. On the one hand, individuals benefit from being in the labor market since they can have a higher labor market income than while studying, they avoid the disutility of being in school and they do not have to pay a tuition fee or other costs that are related to schooling. On the other hand, there are also benefits to staying in school. These benefits consist of increased future earnings, not having to pay a fine if leaving school before the minimum school leaving age and the value that the individual gets from additional skill investment.

The government can affect the individual's schooling choice by supply-side reforms. Since the optimal schooling decision for the individual depends on the fine they need to pay if they drop out of school before the minimum school leaving age, governments can encourage additional schooling by increasing either the fine or the minimum school leaving age. A supply-side reform increasing the minimum school leaving age is seen as exogenous from the individual's perspective and increases the benefits of staying in school since it increases the period over which the individual is exposed to the fine. The effect of increasing the minimum school leaving age on the schooling decision acts through various effects: wealth effects, effects due to an increased marginal value of skill, effects through a higher stock of skills, health effects and an increased life length. All of these effects have a positive impact on the schooling decision, except the wealth effect which can be either positive or negative.

Not all individuals are affected by an increase in the minimum school leaving age. These are the individuals that even in absence of the reform would have chosen to leave school long before reaching the minimum school leaving age or completed many more years of education than what is compulsory. The marginal individual is however affected by the reform in her schooling decision. The marginal individual stays in school due to the fine she has to pay if leaving school before reaching the minimum school leaving age, but would otherwise prefer to enter the labor market to increase her income. The marginal individual thus completes the required years of schooling. If the minimum school leaving age is increased and the fine for dropping out of school before that age is set high enough, then the individual on the margin will stay in school until she reaches the new school leaving age. The individuals, whose schooling decision is affected in this way by an increase in the minimum school leaving age are called compliers. However, even though a complier increases years of completed education, her health, health behavior and longevity will only be increased if the increased years of schooling increase her wealth and skills. The size of these effects will depend on institutions as well as economic and social conditions (Galama et al., 2018).

## 3.3 The effect of compulsory school reforms on health behavior

In the theoretical framework by Galama et al. (2018), compulsory schooling reforms have an effect on health behaviors among compliers. The consumption decision of unhealthy and healthy goods depends on the marginal cost and marginal benefit of the consumption. The marginal benefit of consuming an unhealthy good is the discounted marginal utility from the consumption. The marginal cost of consuming an unhealthy good is the monetary cost and the health cost (i.e. the product of the amount of health lost due to consuming the good and the relative marginal value of health).

An increase in the minimum school leaving age can either have a positive or a negative effect on the consumption of unhealthy goods among the compliers. Such a supply-side reform will affect the consumption of unhealthy goods in two ways. On the one hand, it will have a wealth effect since increased years of schooling will increase wealth. This will enable more consumption of unhealthy goods. On the other hand, when wealth increases due to the increase in years of schooling, this will increase the marginal value of health in relation to wealth. The reasoning behind this is that when an individual gets more wealthy, she will value health higher in relation to consumption as a consequence of health increasing the length of life. This increase in the relative marginal value of health causes the cost of consumption of unhealthy goods to increase as well. This is called the health cost effect. Since the wealth effect and the health cost effect goes in different directions, it is unclear whether increasing the minimum school leaving age has a positive or negative net effect on unhealthy consumption among compliers (Galama et al., 2018). Since the magnitude of the health cost effect increases with the amount of health that is lost due to unhealthy consumption, the net effect is likely to be positive for goods that are moderately unhealthy but negative for goods that are severely unhealthy (Van Kippersluis and Galama, 2014). Galama et al. (2018) note that this assumes perfect information regarding health effects.

The effect of increased compulsory schooling on the consumption of healthy goods among the compliers is more clear. The wealth effect and the health cost effect are no longer ambiguous. This is because the increased wealth due to more years of schooling will both enable more consumption of healthy goods and encourage more healthy behavior due to an increased relative marginal value of health. In this case, the net effect of increasing the minimum school leaving age thus has a positive effect on consumption. It is, therefore, possible that the effect of increasing years of compulsory schooling would have larger effects on the consumption of healthy goods than the consumption of unhealthy goods. In addition, if this theoretical framework is applied to health behaviors such as engaging in physical activity or wearing a face mask and washing hands more often during the Covid-19 pandemic in order to avoid contamination, there is no cost to more a healthy behavior. In these cases, there is no direct wealth effect. Instead, the net effect of increasing the minimum school starting age should equal the health cost effect. By evaluating the effect of different health behaviors, we could thus potentially say something about the relative importance of the wealth effect and the health cost effect.

# 4 Data

## 4.1 SHARE data

This paper uses data from the *Survey of Health, Ageing and Retirement in Europe* (SHARE). The survey is conducted among people aged 50 or older in 28 European countries and Israel and covers the topics of public health and socioeconomic living conditions. SHARE provides longitudinal information at an individual level and is the largest European panel study with microdata on health and socioeconomic variables. The first survey was conducted in 2004 (SHARE, 2022a). I use data from waves 1-2 and 4-8 in this paper (Börsch-Supan, 2022a-2022g). Wave 3 was a survey on life history information and will therefore not be used (SHARE 2022b). This paper uses data collected between 2004 when the collection of information for the first wave was started, and 2020, when the data collection for wave 8 was finished (SHARE, 2022a). Since the SHARE questionnaires are similar across waves and individuals participate in more than one wave, many individuals have responded to the same question multiple times. In these cases, I use the answer given in the latest wave in which the individual participated. Consequently, I use a cross-sectional dataset. Some information, for example concerning education and immigration, is given in a baseline questionnaire that the individual only answers the first time she participates in SHARE. Concerning this information, I use data from the first wave in which the individual participated. The sample constructed using this data is referred to as the main sample.

In addition to the main sample, I construct a sample with data on individuals' health behaviors during the Covid-19 pandemic. During the data collection for wave 8, the Covid-19 pandemic hit Europe. As a response to this, SHARE constructed a specific Corona questionnaire in addition to the regular SHARE questionnaire. The first SHARE Corona survey was conducted in 2020 and the second in 2021. The corona surveys cover topics related to health and socioeconomic living conditions and include for example information about the individual's health behaviors in relation to the pandemic (SHARE, 2022a). In addition to the regular waves, I will use the cross-sectional data collected in these two Corona surveys to study health behaviors in an *at-risk* situation (Börsch-Supan, 2022h, 2022i). The sample constructed using this data from the Corona surveys will be referred to as the Covid-19 sample.

#### 4.1.1 Years of education

To measure education, I use the number of years that the individual has spent in education. The information about the individual's years of education is given in the baseline questionnaire. In wave 1 the information on years of education in the SHARE data is calculated from the individual's highest ISCED level. In the following waves, the individuals were asked about the number of years they spent in education (SHARE, 2022b). The main sample initially includes 34,093 observations and the Covid-19 sample initially includes 14,366 observations. I exclude individuals who do not know how many years they have spent in education, are still in school or refuse to answer the question. I also exclude the individuals for which the reported number of years of education is suspected to be wrong. In wave 7 there are several outliers. Based on the fact that there are no individuals in the other waves with more than 26 years of education, I exclude those individuals with more than 26 years of schooling in wave 7 to avoid measurement error. After having excluded these individuals, the main sample consists of 33,654 observations and the Covid-19 sample consists of 14,200 observations.

### 4.1.2 Main health behaviors

I consider four main health behaviors in this paper: smoking, engagement in physical activity, alcohol consumption, and consumption of fruits and vegetables. These health behaviors are studied using the main sample. Smoking is a binary variable taking the value of one if the individual has ever smoked daily. The individual's engagement in physical activity is a binary variable taking the value of one if the individual engages in sports or activities that are vigorous at least once per week. The information on smoking and physical activity is collected from waves 1-2 and waves 4-8. The consumption of alcohol is measured as the number of units of alcoholic beverages consumed in the last seven days. This information is taken from waves 6-8. The consumption of fruits and vegetables is measured using a binary variable taking the value of one if the individual, in a regular week, consumes fruits or vegetables at least three times per week. This information is collected from waves 4-8.

### 4.1.3 Covid-19 related health behaviors

In addition to the main health behaviors, I include three Covid-19 related health behaviors: variables for using a facemask, washing hands more often than usual, and wanting to get vaccinated against Covid-19. These health behaviors are studied using the Covid-19 sample. The information about the first two variables is collected from the first Corona survey. In this questionnaire, the individuals were asked how often they wore a face mask when they went outside their home to a public space. The variable indicating the use of a facemask is a binary variable taking the value one if the individual reported "always" or "often" wearing a face mask and zero if the individual reported "sometimes" or "never" wearing a face mask. The variable for washing hands is also a binary variable. The value for this variable is one if the individual reported washing hands more often than before the onset of the pandemic and zero if the individual had not changed behavior in this regard. The variable indicating a positive attitude towards the Covid-19 vaccine is created from information in the second Corona survey. The variable is binary and takes the value one if the individual, at the time of the interview, was already vaccinated, had a vaccination scheduled, or wanted to get vaccinated. The variable takes the value of zero if the individual did not want to get vaccinated or was still undecided at the time of the interview.

## 4.2 Compulsory schooling reforms

Hofmarcher's (2021) database with information on compulsory schooling reforms in European countries during the 20th century is used to construct the instrumental variable for years of education. The database contains all compulsory schooling reforms in the included countries for birth cohorts between 1932 and 1995 at ISCED level 1 or above (Hofmarcher, 2021). This means that the database includes compulsory schooling reforms from the primary education level or above (OECD, 2015). Both reforms that decreased and increased years of compulsory schooling are included. Hofmarcher's database includes information about the year the reform was enacted, the first birth cohort that was affected by the reform, as well as the school starting age and the number of years of compulsory schooling before and after the reform (Hofmarcher, 2021).

To ensure that the number of compulsory years of schooling applies to all individuals in the sample, I have to exclude the individuals that immigrated to the country in which the survey was conducted when the individual was older than the school starting age. This is done by combining information from the SHARE data and the compulsory school reform data. Using the information from the SHARE baseline questionnaire on whether the individual was born in the country of the interview, year of birth and year of immigration to the country of the interview, I construct a variable for immigrants that indicate the age when he or she immigrated to the country. Combining this with the information on school starting age, I exclude individuals that immigrated to the country of the interview at an older age than the school starting age that applied to the individual's birth cohort in the country of immigration. After having excluded these individuals the main sample consists of 31,777 observations and the Covid-19 sample consists of 13,726 observations.

## 4.3 Descriptive statistics

Table 1 presents descriptive statistics of the main sample and the Covid-19 sample. It includes information on education variables, socioeconomic characteristics and health behaviors. The Covid-19 related health behaviors are only presented for the Covid-19 sample since only individuals in this sample answered questions about these behaviors. The samples include individuals up to five birth cohorts before and after a reform. Since the main sample includes individuals from multiple waves of SHARE, while the Covid-19 sample only includes individuals participating in the Corona surveys, the Covid-19 sample is substantially smaller than the main sample. The main sample includes 31,777 individuals and the Covid-19 sample includes 13,726 individuals.

Descriptive statistics are similar across the samples. This suggests that even though the Covid-19 sample is relatively small, it is representative of the SHARE population. The mean years of education in the main sample is approximately 11 and ranges from 0 to 25 in both samples. The mean years of compulsory education is around 7.7 in the main sample and 7.6 in the Covid-19 sample, ranging from 3 to 11 in both samples. The two samples are also similar in mean age, the proportion of men and women, as well as the share of individuals born in the country of interview. Interestingly, the samples are very similar considering the health behaviors of drinking, smoking, engagement in physical activity and consumption of fruits and vegetables. This is reassuring and suggests that the results from the different samples can be compared. Worth noting is the large share of individuals that report eating vegetables or fruits at least three times per week. The share of 90 percent indicates small variation across individuals for this variable. The same pattern is noted for all Covid-19 related health behaviors in the Covid-19 sample. The share of individuals in the sample that report having adopted these health behaviors ranges from 83 to 89 percent, indicating that the variation in health behavior is relatively small for these variables.

Table 1: Descriptive statistics

| | **Main sample** | | | | | **Covid-19 sample** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs | Mean | Std. Dev. | Min | Max | Obs | Mean | Std. Dev. | Min | Max |
| *Education* | | | | | | | | | | |
| Years of compulsory education | 31,777 | 7.736 | 1.476 | 3 | 11 | 13,726 | 7.644 | 1.626 | 3 | 11 |
| Years of education | 31,777 | 11.217 | 3.975 | 0 | 25 | 13,726 | 11.261 | 4.059 | 0 | 25 |
| *Socioeconomic characteristics* | | | | | | | | | | |
| Age | 31,777 | 65.504 | 8.516 | 50 | 99 | 13,726 | 66.454 | 7.437 | 50 | 99 |
| Male | 31,777 | 0.451 | 0.498 | 0 | 1 | 13,726 | 0.423 | 0.494 | 0 | 1 |
| Year of birth | 31,777 | 1949.4 | 8.273 | 1918 | 1968 | 13,726 | 1951.1 | 7.015 | 1921 | 1968 |
| Born in the country of interview | 31,777 | 0.992 | 0.087 | 0 | 1 | 13,726 | 0.993 | 0.082 | 0 | 1 |
| In good subjective health | 31,154 | 0.621 | 0.485 | 0 | 1 | 13,712 | 0.645 | 0.478 | 0 | 1 |
| *Health behaviors* | | | | | | | | | | |
| Units of alcoholic beverage last 7 days | 19,774 | 3.938 | 7.422 | 0 | 140 | 12,395 | 3.801 | 7.355 | 0 | 130 |
| Ever smoked daily | 18,582 | 0.452 | 0.498 | 0 | 1 | 9,862 | 0.417 | 0.493 | 0 | 1 |
| Vigorous physical activity (at least once/week) | 28,786 | 0.462 | 0.499 | 0 | 1 | 12,896 | 0.475 | 0.499 | 0 | 1 |
| Eat fruits/vegetables (at least three times/week) | 18,498 | 0.904 | 0.295 | 0 | 1 | 10,141 | 0.905 | 0.293 | 0 | 1 |
| *Covid-19 related health behaviors* | | | | | | | | | | |
| Wear facemask often | - | - | - | - | - | 11,558 | 0.829 | 0.376 | 0 | 1 |
| Wash hands more often than before the pandemic | - | - | - | - | - | 13,494 | 0.890 | 0.314 | 0 | 1 |
| Want to get vaccinated against covid-19 | - | - | - | - | - | 11,501 | 0.847 | 0.360 | 0 | 1 |

Note: The samples include up to five birth cohorts before and after a reform.

# 5 Empirical approach

## 5.1 Empirical specification

In estimating the effect of education on health behaviors, there is a risk of biased results due to endogeneity problems that stem from omitted variable bias (OVB) (Angrist and Pischke, 2009). If education is correlated with some variable that is also correlated with health behaviors, this will lead to biased results. This would for example be the case if individuals with better health or specific health behaviors select into higher education. Since there might be some unobserved variables that are correlated with both health behaviors and education, estimating a simple OLS regression implies a risk of biased results even if controls are included. In order to avoid endogeneity problems when estimating the effect of education on health behaviors, I follow Hofmarcher (2021) and Brunello et al. (2013a) and construct an IV for education, exploiting exogenous variation in years of compulsory schooling across and within European countries. The variation in compulsory schooling stems from compulsory schooling reforms that were implemented in different

European countries in different time periods after the Second World War. In line with a DiD approach, the IV thus takes advantage of exogenous variation in compulsory schooling across birth cohorts and countries (Hofmarcher, 2021). Using this IV allows me to estimate the causal effect of education on a series of health behaviors. I estimate the following equations to obtain my IV estimate:

$$Y_{ibc} = \alpha_0 + \alpha_1 \widehat{Education}_{ibc} + \alpha_2 X_i + \gamma_c + \delta_b + \epsilon_{ibc} \tag{1}$$

$$Education_{ibc} = \beta_0 + \beta_1 CompEducation_{ibc} + \beta_2 X_i + \mu_c + \theta_b + \upsilon_{ibc} \tag{2}$$

Equation (1) represents the second stage equation. This equation is used to estimate the effect of education on the different health behaviors. In this equation $Y_{ibc}$ is the health behavior for individual $i$, in birth cohort $b$ and country $c$. The coefficient $\alpha_1$ captures the effect of education on the health behavior studied. Instead of using the number of years of schooling, I use an instrument for education that is obtained by running regression (2). Equation (2) represents the first stage regression. This regression estimates $Education_{ibc}$, representing the number of years of schooling for individual $i$ in birth cohort $b$ and country $c$. $CompEducation_{ibc}$ represents the number of years of compulsory schooling that applied to individual $i$ in birth cohort $b$ and country $c$. $\beta_1$ thus captures the effect of the number of compulsory school years on the total number of years in education. $X_i$ is a vector of control variables that include the individual's age at the time of the interview. When running the regressions for the main sample, I also include wave in the vector of controls. $\gamma_c$ and $\mu_c$ are country fixed effects. $\delta_b$ and $\theta_b$ are birth cohort fixed effects. The fixed effects allow to control for differences across countries and birth cohorts. $\epsilon_{ibc}$ and $\upsilon_{ibc}$ are error terms. Following Hofmarcher (2021), standard errors are clustered at country-by-birth-cohort level.

In order to be able to evaluate the results obtained by the IV regression, I also estimate the relationship between education and health behaviors by running a simple OLS regression. Even though the results from this OLS regression are likely to be biased, they will be used as baseline estimates to which the IV estimates will be compared. The simple OLS regression is obtained by estimating the following equation:

$$Y_{ibc} = \alpha_0 + \alpha_1 Education_{ibc} + \alpha_2 X_i + \gamma_c + \delta_b + \epsilon_{ibc} \tag{3}$$

Equation (3) represents the structural form equation. It is similar to equation (2) but differs in the variable for education. Instead of using the instrument for years in education, $Education_{ibc}$ represents the true number of years that individual $i$ in birth cohort $b$ and country $c$ spent in education. $\alpha_1$, therefore, captures the effect of education on the health behavior, but is likely to be biased due to OVB. As in equation (1), $X_i$ is a vector of controls, including age for both samples and wave for the main sample. $\gamma_c$ and $\delta_b$ are country and birth cohort fixed effects.

I only include individuals within a limited sample window of ten years for each reform. This is done in line with similar studies that use compulsory schooling reforms to construct an IV for education (see for example Hofmarcher, 2021; Brunello et al., 2013a). I thus use a five-year bandwidth to construct the sample. For each reform I include the last five birth cohorts not affected by the reform and the first five birth cohorts affected by the reform in the sample. This resembles a RD design and is used in order to limit bias in the results. Since many countries experienced several compulsory schooling reforms during the period studied, some of which followed upon each other within only a few years, using a bandwidth ensures that the individuals are only affected by the specific reform and not by other compulsory schooling reforms. However, in some countries, two compulsory schooling reforms were implemented even within the bandwidth of five years. In these cases, I only include the first reform and the post-treatment period is shortened to avoid individuals being affected by another compulsory schooling reform in their educational decisions. This is done to create clear pre-treatment and post-treatment groups.

## 5.2 The linear probability model

I estimate regressions using a linear probability model (LPM) and all dependent variables are limited dependent variables (LDV). All health behaviors considered, except the variable measuring alcoholic consumption, are binary variables. While the variable measuring alcoholic consumption is not binary, it is still an LDV since it can only take positive values. The risk of using LPM when estimating regressions with LDVs is that the limits of the dependent variables are not respected (Angrist and Pischke, 2009). Consequently, the fitted values for the binary variables risks being negative or larger than one and the fitted values for the non-negative variable of alcohol consumption risk being negative. Despite this, Angrist and Pischke (2009) argue that LPM can still be used to estimate causal effects. The reason for this is that even though a nonlinear model, such as the probit model for example, might fit the conditional expectation function of an LDV better since it respects the boundaries of the LVD, the difference between using such a model and the LPM is small when looking at marginal effects. Similar to Hofmarcher (2021), I therefore run an LPM to estimate the marginal effect of education on the series of health behaviors. Using an LPM implies that the standard errors are heteroscedastic and robust standard errors are therefore used when running the regressions (Angrist and Pischke, 2009).

## 5.3 The IV-approach and the LATE estimate

### 5.3.1 Assumptions

The base for using an IV approach is that the instrument should be correlated with the causal variable of interest but uncorrelated with other observed or unobserved variables that can explain the outcome variable. Four assumptions need to be fulfilled for the instrument to be valid. First, it needs to exist a first stage (Angrist and Pischke, 2009). In this setting, this means that the compulsory schooling reforms should have a significant effect on years spent in education. The number of compulsory years of schooling is likely to be correlated with the total number of years in education, as predicted in the theoretical framework developed by Galama et al. (2018). The existence of a first stage is examined in section 7.1 and the results indicate that there exists a strong first stage.

Second, the independence assumption needs to be fulfilled. This assumption requires that the instrument is as good as randomly assigned (Angrist and Pischke, 2009). In this setting, it implies that the number of years of compulsory schooling assigned to an individual is independent of her health behavior and her total number of years of schooling. This assumption is likely to be fulfilled since I use exogenous variation in compulsory years of schooling across birth cohorts and countries, only include individuals within the bandwidth of five years around each reform and control for birth cohort and country fixed effects. The number of compulsory years of schooling that apply to an individual in the sample is thus likely to be as good as random. The independence assumption is sufficient for interpreting the effect of the instrument on the outcome variable in the second stage as causal. In addition, it ensures that the effect of compulsory schooling on the total years in education in the first stage is causal (Angrist and Pischke, 2009).

Third, the exclusion restriction needs to be fulfilled (Angrist and Pischke, 2009). This implies that changes in compulsory schooling should only affect the health behaviors through the number of years of schooling. This assumption is violated if for example socioeconomic living conditions or health during childhood is correlated with the probability of being affected by a compulsory schooling reform. As Hofmarcher (2021) states, this is not likely to be the case since only the individual's year of birth decides whether the individual was affected by the reform or not. Compulsory schooling has been free of charge in all the included countries during the period studied, which reduces the risk of a potential impact of socioeconomic living conditions (Hofmarcher, 2021). Hofmarcher also underlines that since all reforms were enacted after the birth of the first cohort affected, parents could not time the birth of their children in order for them to be either affected or unaffected by a compulsory schooling reform. Potential differences in probabilities of being affected by a reform that stems from socioeconomic or health differences on a country-year level should be controlled for by the use of country fixed effects. Such difference would for example appear if the timing of the reforms depends on the wealth level or welfare level of the country. Another potential concern for the exclusion restriction is if the compulsory schooling reforms are correlated with other changes in education so that other factors than the number of compulsory schooling years affect health behaviors. If the compulsory schooling reforms correlate with changes in school quality or in the curriculum (Hofmarcher, 2021), so that the students for example learn more about different health behaviors, the exclusion restriction is threatened. Several studies (see for example Brunello et al., 2013b) examine the effect of compulsory schooling reforms in Europe on school quality and do not reject the validity of the instrument for education created using compulsory schooling reforms. Overall, the exclusion restriction seems to be fulfilled in this setting.

A fourth assumption is the monotonicity assumption, stating that the individuals that are affected by the instrument need to be affected in the same direction (Angrist and Pischke, 2009). This indicates that for the instrument to be valid, there cannot be any defiers. That means that there cannot be any individuals that decrease their total years of schooling when the years of compulsory schooling increase and increases their total years of schooling when the years of compulsory schooling decrease. According to the theoretical model by Galama et al. (2018), this kind of behavior is irrational. Moreover, a potential problem with defiers is handled in section 6, where I exclude reforms for which compulsory schooling shows a strong significant negative effect on total years in education from the analysis.

In conclusion, exploiting the exogenous variation in compulsory schooling across birth cohorts and countries, the instrument is unlikely to be correlated with other determinants of education. Using this instrument for education thus solves the endogeneity problem by making the causal variable of interest uncorrelated with the error term.

#### 5.3.2 The LATE estimate

Following the theoretical model by Galama et al. (2018), it is likely that there exist heterogeneous treatment effects. This means that not all individuals are affected by the instrument in the same way (Angrist and Pischke, 2009). In this setting, this implies that the compulsory schooling reforms affect individuals' total number of years in education in different ways. The model by Galama et al. (2018) predicts that only individuals at the margin change their total years of education due to the implementation of a compulsory schooling reform. These are the individuals for which, in the absence of the reform, the number of compulsory years of schooling equals the number of total years of schooling. This group of individuals, compliers, thus increase their total years of schooling due to the instrument. Since changes in reforms will not affect individuals who would have chosen considerably less education, never-takers, or individuals who would have chosen considerably more education, always-takers, in the absence of the compulsory school reform, these individuals will not be accounted for in the instrument. The fourth group of individuals is defiers (Angrist and Pischke, 2009). These individuals acting in the opposite direction to the compulsory schooling reform are excluded from the analysis.

Due to these heterogeneous treatment effects, given that the monotonicity assumption holds, the estimated effect is the Local Average Treatment Effect (LATE). This indicates that the estimated effect is the effect among the compliers (Angrist and Pischke, 2009). Since, according to the theoretical model by Galama et al. (2018), the compliers are individuals for which the number of compulsory years of schooling is close to the total years in education, the LATE applies to individuals with low levels of education. Since it is possible that this group of individuals is affected differently by increased education than always-takers or never-takers, this decreases the external validity of the results.

# 6 Selection of reforms

All of the reforms included in the Hofmarcher (2021) database cannot be used in the analysis. The original database covers 73 reforms in countries represented in the SHARE data. I base the selection of reforms on the main sample. Since that sample only covers birth cohorts between 1901 and 1970 I only include reforms that became effective for individuals born in 1968 at the latest, in order to have at least three post-periods for each reform. The first birth cohort affected by a reform in the database is individuals born in 1923. I, therefore, do not need to exclude any early reforms due to too few pre-periods. I disregard reforms that have none or too few observations in the SHARE data for the relevant birth cohorts. Some countries experienced the implementation of several compulsory schooling reforms within the five-year bandwidth. Consequently, some post-periods coincide with the post-period of a previous reform. In these cases, I exclude the latest reform in order to create clear pre-treatment and post-treatment periods of at least three years for each reform.

When this simple first selection has been done, 37 compulsory schooling reforms remain. Table A1 in the appendix lists these reforms. I examine the effect of each of these reforms on years of education using an RD design. This testing of the existence of a first stage for all reforms is done in line with Hofmarcher (2021) and Brunello et al. (2013a). For each reform, I create a sample window of five years before and five years after the implementation of the compulsory schooling reform. In cases when one reform was followed by another within a five-year period, the post-treatment period is shortened to avoid that individuals are affected by another schooling reform in their educational decision. For each reform, I plot the first stage of mean years of education for the birth cohorts in the sample window. This is done in order to visually examine if there is a discontinuity at the threshold of the implementation of the reform. Figure A1 in the appendix presents the visual presentation of the first stage by reform. In addition to the visual presentation, I run the following regression for each reform for the individuals within the sample window:

$$Education_i = \alpha + \beta CompEducation_i + \gamma Trend_i + \epsilon_i \tag{4}$$

In Equation (4), $Education_i$ represents the total years of schooling for individual $i$, $CompEducation_i$ is the years of compulsory schooling for individual $i$, $Trend_i$ is the birth cohort for individual $i$, representing the running variable, and $\epsilon_i$ is an error term. Table A.2 in the appendix presents the regression results for each reform. In line with Hofmarcher (2021), I use robust standard errors and select the reforms that have a t-statistic of the coefficient that is not significantly negative at the 50 percent significance level (i.e. the t-statistic is greater than -0.674). After this selection process is done, 30 reforms remain to be used in the analysis. Table 2 presents all reforms used in the empirical analysis.

Table 2: Selected compulsory schooling reforms

| Country | Years of compulsory schooling | School starting age | School leaving age | Reform date | First cohort affected |
|---|---|---|---|---|---|
| Austria | 8 to 9 | 6 | 14 to 15 | 1962 | 1952 |
| Bulgaria 1 | 7 to 8 | 7 | 14 to 15 | 1959 | 1946 |
| Bulgaria 2 | 8 to 9 | 7 | 15 to 16 | 1969 | 1958 |
| Czechia 1 | 8 to 9 | 6 | 14 to 15 | 1948 | 1934 |
| Czechia 2 | 9 to 8 | 6 | 15 to 14 | 1953 | 1939 |
| Czechia 3 | 8 to 9 | 6 | 14 to 15 | 1960 | 1947 |
| Denmark 1 | 7 to 9 | 7 | 14 to 16 | 1972 | 1958 |
| France | 8 to 10 | 6 | 14 to 16 | 1959 | 1953 |
| Germany | 8 to 9 | 6 | 14 to 15 | 1967 | 1953 |
| Greece | 6 to 9 | 6 | 12 to 15 | 1976 | 1964 |
| Italy 1 | 5 to 8 | 6 | 11 to 14 | 1962 | 1952 |
| Lithuania 1 | 7 to 8 | 7 | 14 to 15 | 1958 | 1945 |
| Lithuania 2 | 8 to 11 | 7 | 15 to 18 | 1970 | 1955 |
| Luxembourg 1 | 7 to 8 | 6 | 13 to 14 | 1945 | 1932 |
| Malta 1 | 8 to 10 | 6 | 14 to 16 | 1974 | 1960 |
| Netherlands 1 | 7 to 8 | 6 | 13 to 14 | 1947 | 1937 |
| Poland 1 | 7 to 8 | 7 | 14 to 15 | 1961 | 1952 |
| Portugal 1 | 3 to 4 | 7 | 12 | 1956/1960* | 1950 |
| Portugal 2 | 4 to 6 | 7 | 12 to 14 | 1964 | 1957 |
| Romania 1 | 7 to 4 | 7 | 11 | 1948 | 1935 |
| Romania 2 | 4 to 7 | 7 | 14 | 1958 | 1947 |
| Slovakia 1 | 8 to 9 | 6 | 14 to 15 | 1948 | 1934 |
| Slovakia 2 | 9 to 8 | 6 | 15 to 14 | 1953 | 1939 |
| Slovakia 3 | 8 to 9 | 6 | 14 to 15 | 1960 | 1947 |
| Slovakia 4 | 9 to 10 | 6 | 15 to 16 | 1978 | 1964 |
| Slovenia 1 | 4 to 7 | 7 | 11 to 14 | 1945 | 1935 |
| Spain 1 | 7 to 8 | 6 | 13 to 14 | 1970 | 1958 |
| Sweden 1 | 6 to 7 | 7 | 13 to 14 | 1936 | 1923 |
| Sweden 2 | 7 to 9 | 7 | 14 to 16 | 1962 | 1951 |

Note: The information on the compulsory schooling reforms is collected from the database of reforms in Hofmarcher (2021).
*The reform was enacted for boys in 1956 and for girls in 1960.

# 7 Results

## 7.1 First stage results

### 7.1.1 Visual presentation of the first stage

In order to examine the existence of a first stage, I plot the mean years of education by distance from the first birth cohort affected by the reform. Within the five-year bandwidth, included birth cohorts from each reform are bunched together with the birth cohorts from the other reforms based on their distance to the first birth cohort affected by the reform. The reforms are divided into two groups: reforms that increased the number of years of compulsory education and reforms that decreased the years of compulsory education. While the first group includes 27 reforms, the latter only includes three reforms. The existence of a first stage is examined separately for each group of reforms and for both samples.

Figures 1 and 2 show the results for the main sample. Figure 1 plots the mean years of education by distance from the first birth cohort affected by a reform that increased the number of years of compulsory schooling. There is a discontinuity at the threshold of the implementation of the reform, suggesting that increasing the years of compulsory education on average seems to increase the total years of schooling. The slopes of the fitted lines before and after the threshold are both positive, indicating an increase in total education over time. Figure 2 plots the mean years of education by distance from the first birth cohort affected by a compulsory schooling reform that decreased the number of compulsory years of schooling. Figure 2 also shows a discontinuity at the threshold, but in the opposite direction than that in Figure 1. This is reassuring since it suggests that, in line with theory, decreasing the number of years of schooling seems to decrease the total number of years of schooling for the average individual in the main sample.

Comparing Figure 1 and Figure 2, it is apparent that the fitted lines in Figure 1 follow the mean values for years of education more closely than the fitted lines in Figure 2, where the means for the birth cohorts after the reforms do not seem to follow a linear trend. This is probably due to the fact that there are only three selected reforms that decreased the number of years of compulsory schooling. It is reassuring that the fitted lines on both sides of the threshold have the same sign of the slope. Since this graph only includes three reforms, it is unlikely to have a large impact on the first stage. This visual presentation suggests that there exists a first stage for the main sample.

Figure 1: Main sample - Increasing compulsory education

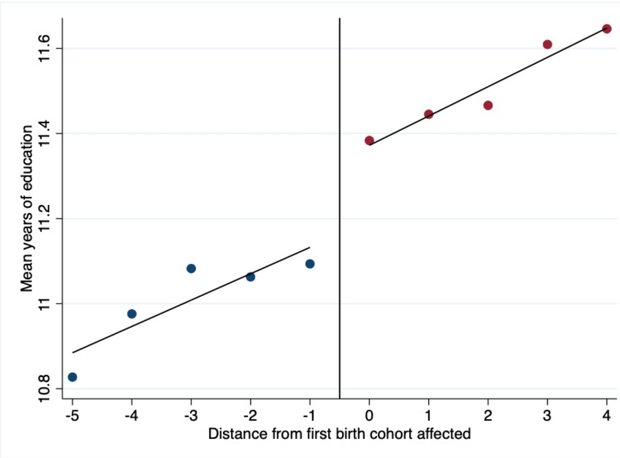
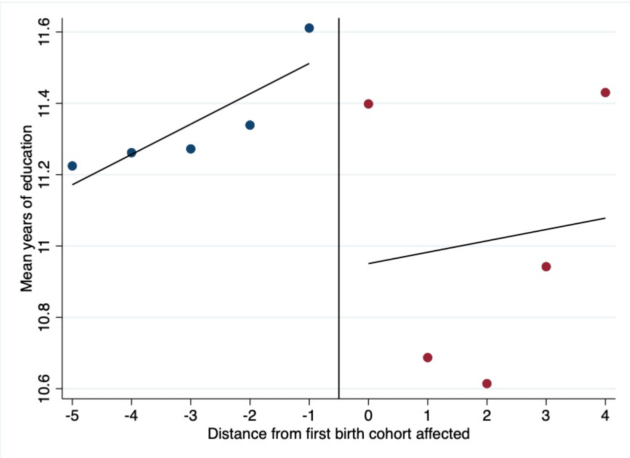



Figure 2: Main sample - Decreasing compulsory education




Note: Mean years of education by distance from the first birth cohort affected by a reform for the main sample. Figure 1 includes all reforms increasing years of compulsory education. Figure 2 includes all reforms decreasing years of compulsory education.

Figures 3 and 4 plot the mean years of schooling by distance from the first birth cohort affected for the Covid-19 sample. Similar to what was seen for the main sample, there seems to be a positive jump at the threshold for the reforms increasing the years of compulsory schooling in Figure 3. Even though the Covid-19 sample is smaller than the main sample, it seems to be similar to the main sample regarding the effect of a positive compulsory schooling reform on the mean years of education. The jump at the threshold is about the same size as in the main sample. Similar to the main sample, the positive slope of the fitted lines before and after the threshold indicates a trend of total education increasing over time. The coherency of the first stage across the two samples is reassuring for the validity of the instrument.

Figure 4 plots the mean years of education for birth cohorts around reforms that decreased compulsory education. In line with Figure 3, there is an apparent discontinuity at the threshold. This indicates that a decrease in the years of compulsory education seems to decrease the total number of years in education for the average individual in the Covid-19 sample. Similar to Figure 2, the means in education for the birth cohorts affected by the reforms do not follow the linear trend closely. Moreover, the slope of the fitted line before and after the threshold has different signs. This is in contrast to what was found for the decreasing schooling reforms in the main sample and can be due to the fact that the Covid-19 sample includes fewer observations. This introduces some uncertainty to the strength of the first stage. But as previously mentioned, only three out of 30 selected reforms decreased the number of compulsory years of schooling. It is, therefore, reasonable to assume that the different slopes of the fitted lines in Figure 4 do not have a large effect on the first stage. To conclude, the visual presentation of the first stage in the Covid-19 sample indicates that there exists a first stage, even though it is potentially weaker than in the main sample.

Figure 3: Covid-19 sample - Increasing compulsory education

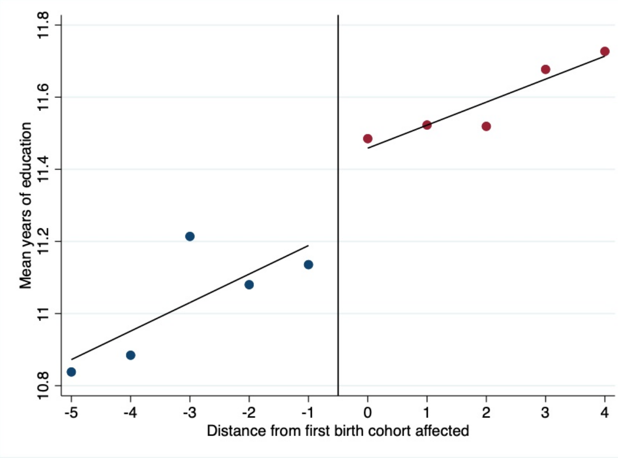
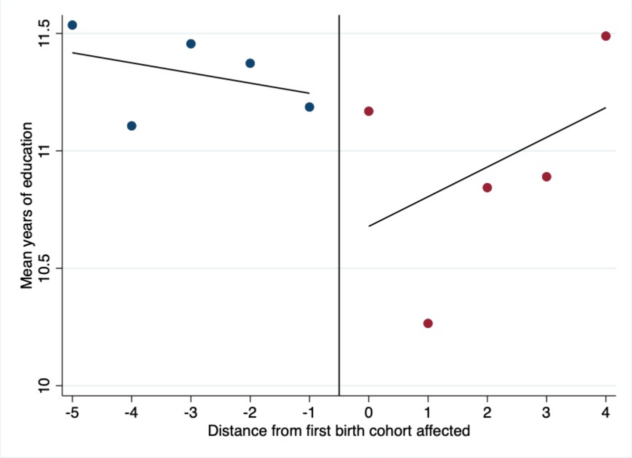



Figure 4: Covid-19 sample - Decreasing compulsory education




Note: Mean years of education by distance from the first birth cohort affected by a reform for the Covid-19 sample. Figure 3 includes all reforms increasing years of compulsory education. Figure 4 includes all reforms decreasing years of compulsory education.

### 7.1.2 Event study results

To further examine the existence of a first stage, I perform an event study analysis. The event study includes both reforms increasing and decreasing the years of compulsory education.[1] I use country and birth cohort fixed effects in the event study and include five leads and four lags. Standard errors are clustered at the country-by-birth-cohort level.

Figure 5 and Figure 6 graphically present the results from the event studies for the main sample and the Covid-19 sample respectively, both showing the change in mean years of education in relation to the last cohort that was not affected by the reform. For both samples, the estimates are close to zero before the implementation of the reform but positive for the birth cohorts affected by the reform. While there is no difference between the birth cohorts unaffected by the reform, there seems to be an increase in total years of schooling after the reform in relation to the last birth cohort affected. The increase in total years of schooling is however more pronounced for the main sample than for the Covid-19 sample. These results are in line with the visual presentation of the first stage, suggesting that there exists a first stage but that it might be weaker for the Covid-19 sample than the main sample, possibly as a result of fewer observations.

[1] Since this analysis requires a time variable that is normalized around the first birth cohort affected, I disregard two of the reforms decreasing the years of compulsory education (reform 2 in Czechia and Slovakia). These reforms became effective too close in time to a previous reform and would confuse the normalized time variable if they were included.

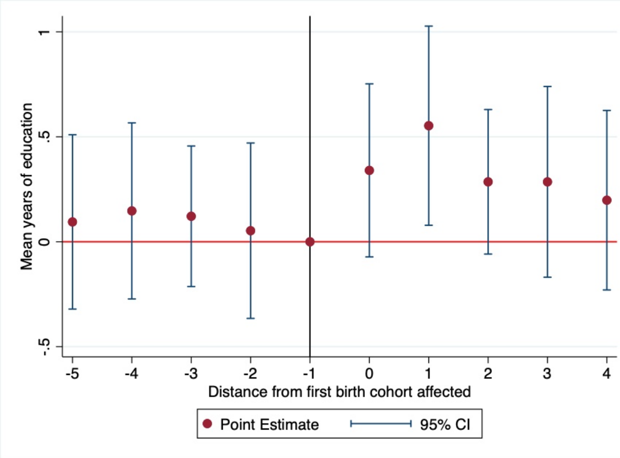



Figure 5: Main sample - Event study

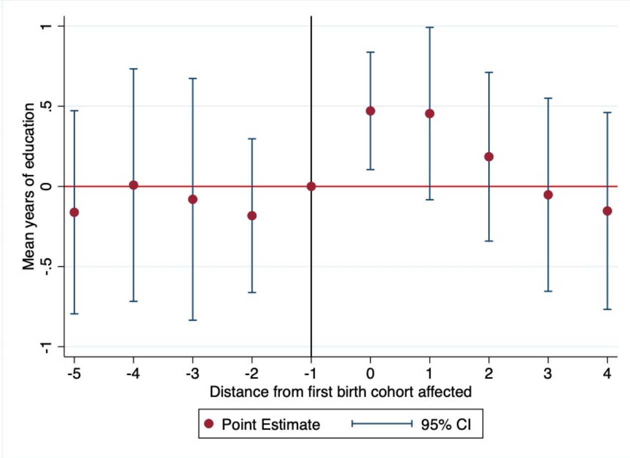



Figure 6: Covid-19 sample - Event study

Note: The vertical line marks the last cohort not affected by the reform. Standard errors are clustered at country-by-birth-cohort level.

## 7.2 Second stage results

Next, I present the second stage results. For each health behavior, I present the results for the OLS regression and the IV regression. A first stage result is also presented for each regression, in which the instrument of compulsory education is regressed on years of schooling.[2] Since I estimate LPM regressions, the coefficients should be interpreted as marginal effects, meaning that they indicate the effect of one additional year of education (Hofmarcher, 2021).

### 7.2.1 Main health behaviors

Columns 1, 4, 7 and 10 of Table 3 report the OLS estimates for smoking, alcohol consumption, physical activity and consumption of fruits and vegetables respectively. These estimates suggest that one additional year of education has a positive effect on all four health outcomes. The estimates are significant for all health outcomes except smoking but are likely to be biased due to omitted variables.

When running the IV regressions, the estimates for smoking, alcohol consumption and consumption of fruits and vegetables change sign and increase in magnitude. Columns 2, 5 and 11 of Table 3 report the results for these variables. One additional year of schooling decreases the probability of having ever smoked by 12.8 percentage points. In comparison to the mean of individuals in the main sample reporting having smoked daily at some point in their lives, this represents a decrease in probability by 28 percent, a considerable effect. The IV estimate for alcohol consumption suggests that one year of

[2]The reason for presenting the first stage for each regression is that the number of individuals for which answers about the different health behaviors are reported varies.

additional schooling decreases the number of drinks during the last seven days by 1.429. This represents a 36 percent decrease. Regarding consumption of fruits and vegetables, the 2SLS estimate indicates that one more year of education decreases the probability of eating fruits or vegetables at least three times per week by 4.63 percentage points. This represents a five percent decrease.

Column 8 of Table 3 reports the IV estimate for the regression on physical activity. In contrast to the other three variables, this estimate stays positive and increases in size, but becomes insignificant when running the IV regression instead of the OLS regression. This suggests that while one more year of schooling seems to be correlated with higher physical activity, there is no evidence of a causal effect of education on physical activity.

Columns 3, 6, 9 and 12 of Table 3 report the results for the first stage regressions. All first stage regressions show that compulsory education has a significant effect on the total years of education. This is in line with the evaluation of the existence of the first stage in section 7.1. The reported estimates suggest that one more year of compulsory education increases the total years in education by 0.16 to 0.17 years. It is reassuring that the first stage estimates are similar in size. The results are further strengthened by the fact that all reported F-statistics of the excluded instrument are larger than 10. According to Staiger and Stock (1997), one needs not to worry about having a weak instrument if the F-statistic is larger than 10. The instrument thus seems to be valid.

### 7.2.2 Covid-19 related health behaviors

Columns 1, 4 and 7 of Table 4 report the estimates from the OLS regressions for use of facemask, washing of hands and wanting to get vaccinated against Covid-19. Similar to the main health behaviors results, the OLS estimates for the Covid-19 related health behaviors are small, positive, and significant. These regressions thus suggest that there is an association between education and the Covid-19 related health behaviors.

When running the IV regression instead of the OLS regression, the magnitude of the effects increases but the estimates become insignificant. Columns 2, 5 and 8 of Table 4 present the 2SLS estimates. While the estimate for the effect of education on the probability of washing hands more often than before the pandemic stays positive, the estimates for the probability of frequently using a facemask and wanting to get vaccinated against Covid-19 change sign. Keeping in mind that the estimates are insignificant, the size of the estimates could give some information about the potential effect of education on the Covid-19 related health behaviors. The estimates suggest that one year of additional education decreases the probability of frequent use of facemask by 1.2 percentage points and the probability of being vaccinated or wanting to get vaccinated against Covid-19 by 3.87 percentage points. While the first estimated effect corresponds to a 1.5 percent decrease, the latter corresponds to a 4.6 percent decrease. The estimated effect of one additional year of education on the probability of washing hands more often than before the pandemic is 2.43 percentage points or 2.7 percent.

The first stage results indicate that one more year of compulsory education significantly increases total years in education by between 0.14 and 0.17 years. Although showing a slightly larger variation than the first stage estimates for the main health behaviors, it is reassuring that all estimates are similar in size. The F-statistics indicate that the instrument is weakly estimated for wanting to get vaccinated against Covid-19, but not for using a facemask or washing hands.

Table 3: Main health behaviors

| | Smoking | | | Alcohol consumption | | | Physical activity | | | Fruit/vegetable consumption | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | OLS | 2SLS | First stage | OLS | 2SLS | First stage | OLS | 2SLS | First stage | OLS | 2SLS | First stage |
| Years of education | 0.000895 | -0.128*** | | 0.0926*** | -1.429** | | 0.00681*** | 0.0301 | | 0.00411*** | -0.0463* | |
| | (0.00132) | (0.0455) | | (0.0159) | (0.601) | | (0.000883) | (0.0319) | | (0.000673) | (0.0250) | |
| Compulsory Education | | | 0.171*** | | | 0.164*** | | | 0.160*** | | | 0.158*** |
| | | | (0.0481) | | | (0.0438) | | | (0.0427) | | | (0.0466) |
| Constant | 1.439*** | | | -4.307 | | | 0.383 | | | 0.927*** | | |
| | (0.484) | | | (6.755) | | | (0.355) | | | (0.268) | | |
| Observations | 18582 | 18582 | 18582 | 19774 | 19774 | 19774 | 28786 | 28786 | 28786 | 18498 | 18498 | 18498 |
| R-squared | 0.060 | -0.897 | | 0.046 | -0.573 | | 0.081 | -0.025 | | 0.057 | -0.394 | |
| F-statistic | 13.07 | 7.202 | 12.61 | 12.88 | 3.298 | 14.01 | 36.16 | 16.56 | 14.01 | 20.70 | 7.539 | 11.49 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions include controls for age and wave. Standard errors are clustered at country-by-birth-cohort level. The sample window is five years before and after each reform.

Table 4: Covid-19 related health behaviors

| | Use of facemask | | | Wash hands more often | | | Vaccinated/want to get vaccinated | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) 2SLS | (3) First stage | (4) OLS | (5) 2SLS | (6) First stage | (7) OLS | (8) 2SLS | (9) First stage |
| Years of Education | 0.00238*** | -0.0122 | | 0.00211*** | 0.0243 | | 0.00659*** | -0.0387 | |
| | (0.000641) | (0.0210) | | (0.000724) | (0.0231) | | (0.000997) | (0.0438) | |
| Compulsory Education | | | 0.168*** | | | 0.165*** | | | 0.143*** |
| | | | (0.0511) | | | (0.0473) | | | (0.0465) |
| Constant | 0.896*** | | | 1.139*** | | | 0.526 | | |
| | (0.281) | | | (0.326) | | | (0.410) | | |
| Observations | 11555 | 11555 | 11555 | 13493 | 13493 | 13493 | 11499 | 11499 | 11499 |
| R-squared | 0.651 | -0.054 | | 0.031 | -0.067 | | 0.244 | -0.265 | |
| F-statistic | 4.615 | 0.272 | 10.73 | 3.374 | 0.933 | 12.17 | 15.04 | 0.461 | 9.48 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions control for age. Standard errors are clustered at country-by-birth-cohort level. The sample window is five years before and after each reform.

## 7.3 Heterogeneous effects

Based on the conclusion of heterogeneous effects of Galama et al. (2018), I split the two samples by gender to get further insight into how education affects health behaviors.

### 7.3.1 Main health behaviors

Table 5 presents the results from running the regressions on the main health behaviors for men and women separately. Columns 1-3 report the estimates for men and columns 4-6 report the estimates for women. The results suggest that the negative effects of education on smoking and alcohol consumption found in the main regression are mainly driven by women. The IV estimate for smoking is significant for both men and women, but the effect is considerably larger for women. While one additional year of schooling decreases the probability of smoking by 7.4 percentage points for men, it decreases the same probability by 16.6 percentage points for women. Compared to the mean of the share of women and men that report having ever smoked daily, this represents a 13 percent decrease for men and a 46.6 percent decrease for women.[3] In contrast, the 2SLS estimate for alcohol consumption is only significant for women. One year of additional education seems to decrease the units of alcohol consumed during the last seven days by 0.96 units for women, corresponding to a decrease of 46.6 percent. Even though the insignificant 2SLS estimate for men is slightly larger, it only corresponds to a 15.7 percent decrease due to the higher mean of alcohol units consumed among men.

No significant effect is found of education on physical activity or consumption of fruits and vegetables when running the IV regressions separately for men and women. While these results strengthen the results for physical activity in the main sample, it is surprising to find no significant effect of education on the consumption of fruits and vegetables when splitting the sample. It is however reassuring that the estimate for both genders is negative and is about the same size as the 2SLS estimate in the main regression.

Three important notes can be made from the first stage results. First, while all first stage estimates report a significant and positive effect of compulsory schooling on total years in education, the effect is larger for men than for women. Second, the reported F-statistics of the excluded instrument are lower than in the main regressions. This is reasonable since the samples are substantially smaller. Third, it is interesting that the instrument only seems to be weak for men in the regression for alcohol consumption, while it is weak for women for smoking, physical activity and consumption of fruits and vegetables. The weak instrument for men when running the regression on alcohol consumption could be a reason for the insignificant result of the effect of education on alcohol consumption. It is

[3]Table A.3 and A.4 in the appendix present descriptive statistics for men and women separately for both samples.

also possible that the weak instrument for women in the regression on fruit and vegetable consumption explains the insignificant result for this estimate.

#### 7.3.2 Covid-19 related health behaviors

Table 6 presents the results from running the regressions for the Covid-19 related health behaviors for men and women separately. Columns 1-3 report the estimates for men and columns 4-6 report the estimates for women. In line with the results of the main regressions, all IV estimates are insignificant. Interestingly, the estimates for men are positive for all health outcomes while the estimates for women are negative for all health outcomes. This indicates that if anything, education can increase protective health behaviors among men, but not among women. The different signs of the estimates for men and women can be a reason for the insignificant result found when running the regressions for both men and women.

Similarly to when examining heterogeneous effects for the main health behaviors, the first stage seems to be significant for both men and women in all regressions, but the effect is stronger for men than for women. Furthermore, the F-statistics reported for the excluded instrument for the IV regressions with Covid-19 related health outcomes are larger for men than for women. The regression for men for wanting to get the vaccine against Covid-19 is however the only regression reporting an F-statistic larger than 10, indicating that the instrument is weak for all other regressions.

Table 5: Heterogenous effects - Main health behaviors

| | **Men** | | | **Women** | | |
|---|---|---|---|---|---|---|
| | **(1)** **OLS** | **(2)** **2SLS** | **(3)** **First stage** | **(4)** **OLS** | **(5)** **2SLS** | **(6)** **First stage** |
| *Panel A: Smoking* | | | | | | |
| Years of education | -0,00951*** | -0,0743* | | 0,00424*** | -0,166** | |
| | (0,00169) | (0,0390) | | (0,00158) | (0,0842) | |
| Compulsory Education | | | 0,231*** | | | 0,138** |
| | | | (0,0710) | | | (0,0540) |
| Constant | 1,119 | | | 1,692*** | | |
| | (0,712) | | | (0,631) | | |
| Observations | 8299 | 8299 | 8299 | 10281 | 10281 | 10281 |
| R-squared | 0,053 | -0,232 | | 0,102 | -1,627 | |
| F-statistic | 19,67 | 7,054 | 10,55 | 4,532 | 2,162 | 6,55 |
| *Panel B: Alcohol consumption* | | | | | | |
| Years of education | -0,0263 | -1,000 | | 0,0951*** | -0,958** | |
| | (0,0254) | (0,826) | | (0,0113) | (0,469) | |
| Compulsory Education | | | 0,178*** | | | 0,169*** |
| | | | (0,0590) | | | (0,0522) |
| Constant | -2,629 | | | -2,675 | | |
| | (11,75) | | | (5,852) | | |
| Observations | 8651 | 8651 | 8651 | 11121 | 11121 | 11121 |
| R-squared | 0,062 | -0,156 | | 0,065 | -0,674 | |
| F-statistic | 3,387 | 2,798 | 9,13 | 24,61 | 1,809 | 10,46 |
| *Panel C: Physical activity* | | | | | | |
| Years of education | 0,00443*** | 0,0229 | | 0,0102*** | 0,0384 | |
| | (0,00119) | (0,0402) | | (0,00140) | (0,0438) | |
| Compulsory Education | | | 0,193*** | | | 0,143*** |
| | | | (0,0577) | | | (0,0510) |
| Constant | 0,116 | | | 1,028*** | 0,406 | -2,333 |
| | (0,511) | | | (0,0474) | (0,712) | (4,434) |
| Observations | 12933 | 12933 | 12933 | 15851 | 15851 | 15851 |
| R-squared | 0,076 | -0,016 | | 0,050 | 0,043 | |
| F-statistic | 12,83 | 9,460 | 11,19 | 115,4 | 231,5 | 7,83 |
| *Panel D: Fruit/vegetable consumption* | | | | | | |
| Years of education | 0,00567*** | -0,0408 | | 0,00404*** | -0,0391 | |
| | (0,000945) | (0,0333) | | (0,000831) | (0,0350) | |
| Compulsory Education | | | 0,212*** | | | 0,133** |
| | | | (0,0658) | | | (0,0541) |
| Constant | 0,990** | | | 0,883** | | |
| | (0,462) | | | (0,349) | | |
| Observations | 8275 | 8275 | 8275 | 10219 | 10219 | 10219 |
| R-squared | 0,063 | -0,281 | | 0,069 | -0,341 | |
| F-statistic | 16,63 | 4,983 | 10,33 | 11,40 | 2,701 | 6,08 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions include controls for age and wave. Standard errors are clustered at country-by-birth-cohort level. The sample window is five years before and after each reform.

Table 6: Heterogenous effects - Covid-19 related health behaviors

| | **Men** | | | **Women** | | |
|---|---|---|---|---|---|---|
| | **(1)** **OLS** | **(2)** **2SLS** | **(3)** **First stage** | **(4)** **OLS** | **(5)** **2SLS** | **(6)** **First stage** |
| *Panel A: Use of facemask* | | | | | | |
| Years of Education | 0.00209** | 0.00201 | | 0.00294*** | -0.0342 | |
| | (0.000980) | (0.0236) | | (0.000867) | (0.0417) | |
| Compulsory Education | | | 0.224*** | | | 0.130** |
| | | | (0.0711) | | | (0.0574) |
| Constant | 0.763* | | | 0.950*** | | |
| | (0.393) | | | (0.339) | | |
| Observations | 5050 | 5050 | 5050 | 6503 | 6503 | 6503 |
| R-squared | 0.629 | 0.001 | | 0.679 | -0.384 | |
| F-statistics | 1.562 | 0.0987 | 9.94 | 4.005 | 0.374 | 5.09 |
| *Panel B: Wash hands more often* | | | | | | |
| Years of Education | 0.00471*** | 0.0518 | | 0.000787 | -0.00225 | |
| | (0.00117) | (0.0346) | | (0.000923) | (0.0325) | |
| Compulsory Education | | | 0.196*** | | | 0.152*** |
| | | | (0.0654) | | | (0.0523) |
| Constant | 0.578 | | | 1.534*** | | |
| | (0.514) | | | (0.429) | | |
| Observations | 5707 | 5707 | 5707 | 7782 | 7782 | 7782 |
| R-squared | 0.035 | -0.291 | | 0.038 | -0.001 | |
| F-statistics | 6.054 | 0.959 | 8.98 | 1.085 | 0.818 | 8.49 |
| *Panel C: Vaccinated/want to get vaccinated* | | | | | | |
| Years of Education | 0.00545*** | 0.00353 | | 0.00691*** | -0.133 | |
| | (0.00126) | (0.0362) | | (0.00123) | (0.108) | |
| Compulsory Education | | | 0.213*** | | | 0.0946* |
| | | | (0.0650) | | | (0.0529) |
| Constant | 1.158* | | | 0.0404 | | |
| | (0.595) | | | (0.538) | | |
| Observations | 4804 | 4804 | 4804 | 6693 | 6693 | 6693 |
| R-squared | 0.230 | 0.004 | | 0.263 | -2.365 | |
| F-statistics | 6.341 | 0.172 | 10.70 | 11.23 | 0.861 | 3.20 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions control for age. Standard errors are clustered at country-by-birth-cohort level. The sample window is five years before and after each reform.

## 7.4 Robustness

In order to test the robustness of the results, I perform a robustness check by varying the bandwidth in line with Hofmarcher (2021). I run the regressions using a four-year bandwidth and a six-year bandwidth, as well as estimating the regression using the full sample, i.e. using no bandwidth. Table 7 presents the results for the robustness check for the main health behaviors and Table 8 presents the results for the robustness check for the Covid-19 related health behaviors.

Running the regressions with a four-year bandwidth and a 6-year bandwidth gives similar estimates for the IV regressions to when using a five-year bandwidth, both for the main health behaviors and the Covid-19 related health behaviors. The estimates for smoking, alcohol consumption and fruit and vegetable consumption are still significant and the estimates for physical activity and the Covid-19 related health behaviors are still insignificant. The sign only changes for the insignificant estimate of the effect of education on the use of facemask when using a four-year bandwidth. The size of the estimates does not change considerably when changing the bandwidth, but when using a six-year bandwidth the estimates for the main health behaviors are somewhat smaller than when using a five-year bandwidth.

The first stage regressions for the main health behaviors and the Covid-19 health behaviors show a similar pattern when varying the bandwidth. The effect of compulsory schooling on total years in education seems to be strongest when using the six-year bandwidth. Using this bandwidth also results in the largest F-statistic for all health behaviors, indicating that the instrument is strongest in this setting. Given the small decrease in estimates for the main health behaviors when using the six-year bandwidth in comparison to the five-year bandwidth, the effect of education on alcohol consumption and smoking might be slightly overestimated. Overall, the results do however not seem to be too sensitive to the choice of bandwidth.

Using the full sample in the analysis of the main health behaviors and the Covid-19 related health behaviors substantially decreases both the size of the estimates and the reported F-statistic for the first stage regressions. This result suggests that using a bandwidth increases the internal validity of the results. It is however reassuring that the negative effect of education on smoking and alcohol consumption is still significant when using the full sample. The negative effect of education on fruit and vegetable consumption becomes insignificant and changes sign. This is in line with the insignificant results when splitting the sample by gender and indicates that the effect of education on fruit and vegetable consumption does not seem to be as robust as that on alcohol consumption and smoking.

Table 7: Robustness test - Main health behaviors

| | 4 years bandwidth | | | 6 years bandwidth | | | Full sample | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) 2SLS | (3) First stage | (4) OLS | (5) 2SLS | (6) First stage | (7) OLS | (8) 2SLS | (9) First stage |
| *Panel A: Smoking* | | | | | | | | | |
| Years of education | 0.00140 | -0.123** | | 0.00125 | -0.0846*** | | -0.00000392 | -0.0466* | |
| | (0.00138) | (0.0506) | | (0.00119) | (0.0293) | | (0.0000265) | (0.0264) | |
| Compulsory Education | | | 0.170*** | | | 0.223*** | | | 0.281** |
| | | | (0.0537) | | | (0.0492) | | | (0.128) |
| Constant | 1.575*** | | | 1.361*** | | | 1.497*** | | |
| | (0.531) | | | (0.449) | | | (0.277) | | |
| Observations | 15624 | 15624 | 15624 | 21353 | 21353 | 21353 | 54085 | 54085 | 54085 |
| R-squared | 0.061 | -0.834 | | 0.056 | -0.399 | | 0.057 | -14.026 | |
| F-statistic | 10.29 | 5.999 | 9.95 | 15.26 | 10.84 | 20.45 | 64.02 | 14.39 | 4.84 |
| *Panel B: Alcohol consumption* | | | | | | | | | |
| Years of education | 0.0875*** | -1.559** | | 0.0915*** | -1.160*** | | 0.00110 | -0.942* | |
| | (0.0179) | (0.707) | | (0.0141) | (0.445) | | (0.000928) | (0.486) | |
| Compulsory Education | | | 0.154*** | | | 0.202*** | | | 0.258** |
| | | | (0.0468) | | | (0.0451) | | | (0.125) |
| Constant | -0.0878 | | | -1.539 | | | -1.103 | | |
| | (7.567) | | | (6.107) | | | (3.965) | | |
| Observations | 16661 | 16661 | 16661 | 22696 | 22696 | 22696 | 55092 | 55092 | 55092 |
| R-squared | 0.045 | -0.648 | | 0.047 | -0.401 | | 0.050 | -28.788 | |
| F-statistic | 8.804 | 2.563 | 10.83 | 15.44 | 3.958 | 20.03 | 6.028 | 2.170 | 4.27 |
| *Panel C: Physical activity* | | | | | | | | | |
| Years of education | 0.00667*** | 0.0147 | | 0.00697*** | 0.0194 | | 0.0000616 | 0.0206 | |
| | (0.000979) | (0.0367) | | (0.000856) | (0.0247) | | (0.0000866) | (0.0178) | |
| Compulsory Education | | | 0.139*** | | | 0.202*** | | | 0.222** |
| | | | (0.0447) | | | (0.0449) | | | (0.0862) |
| Constant | 0.413 | | | 0.294 | | | 0.524** | | |
| | (0.384) | | | (0.325) | | | (0.205) | | |
| Observations | 24229 | 24229 | 24229 | 32966 | 32966 | 32966 | 84749 | 84749 | 84749 |
| R-squared | 0.083 | 0.001 | | 0.081 | -0.004 | | 0.132 | -1.906 | |
| F-statistic | 30.75 | 13.64 | 9.61 | 39.46 | 17.65 | 20.28 | 72.15 | 38.31 | 6.61 |
| *Panel D: Fruit/vegetable consumption* | | | | | | | | | |
| Years of education | 0.00379*** | -0.0463* | | 0.00431*** | -0.0377** | | 0.0000333 | 0.00212 | |
| | (0.000724) | (0.0259) | | (0.000624) | (0.0176) | | (0.0000305) | (0.00652) | |
| Compulsory Education | | | 0.162*** | | | 0.203*** | | | 0.276** |
| | | | (0.0515) | | | (0.0481) | | | (0.127) |
| Constant | 0.809*** | | | 0.805*** | | | 0.869*** | | |
| | (0.293) | | | (0.244) | | | (0.150) | | |
| Observations | 15491 | 15491 | 15491 | 21281 | 21281 | 21281 | 52844 | 52844 | 52844 |
| R-squared | 0.057 | -0.392 | | 0.057 | -0.277 | | 0.046 | -0.096 | |
| F-statistic | 15.77 | 6.535 | 9.93 | 22.08 | 7.055 | 17.76 | 19.98 | 18.29 | 4.71 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions include controls for age and wave. Standard errors are clustered at country-by-birth-cohort level.

Table 8: Robustness test - Covid-19 related health behaviors

| | **4 years bandwidth** | | | **6 years bandwidth** | | | **Full sample bandwidth** | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) 2SLS | (3) First stage | (4) OLS | (5) 2SLS | (6) First stage | (7) OLS | (8) 2SLS | (9) First stage |
| *Panel A: Use of facemask* | | | | | | | | | |
| Years of Education | 0.00272*** | 0.00321 | | 0.00255*** | -0.00509 | | 0.0000147 | 0.00202 | |
| | (0.000700) | (0.0173) | | (0.000587) | (0.0166) | | (0.0000104) | (0.00512) | |
| Compulsory Education | | | 0.197*** | | | 0.205*** | | | 0.493 |
| | | | (0.0545) | | | (0.0515) | | | (0.309) |
| Constant | 1.044*** | | | 0.888*** | | | 0.979*** | | |
| | (0.312) | | | (0.256) | | | (0.191) | | |
| Observations | 9727 | 9727 | 9727 | 13277 | 13277 | 13277 | 29586 | 29586 | 29586 |
| R-squared | 0.646 | 0.002 | | 0.656 | -0.013 | | 0.620 | -0.189 | |
| F-statistic | 5.057 | 0.335 | 13.01 | 6.489 | 0.0700 | 15.86 | 1.673 | 0.947 | 2.54 |
| *Panel B: Wash hands more often* | | | | | | | | | |
| Years of Education | 0.00193** | 0.00691 | | 0.00230*** | 0.0157 | | 0.0000206 | 0.00680 | |
| | (0.000796) | (0.0219) | | (0.000675) | (0.0176) | | (0.0000139) | (0.00724) | |
| Compulsory Education | | | 0.180*** | | | 0.209*** | | | 0.393* |
| | | | (0.0491) | | | (0.0493) | | | (0.212) |
| Constant | 1.040*** | | | 1.167*** | | | 0.775*** | | |
| | (0.365) | | | (0.307) | | | (0.214) | | |
| Observations | 11367 | 11367 | 11367 | 15478 | 15478 | 15478 | 35751 | 35751 | 35751 |
| R-squared | 0.033 | -0.003 | | 0.030 | -0.024 | | 0.028 | -1.061 | |
| F-statistic | 2.413 | 0.539 | 13.43 | 4.517 | 0.935 | 17.94 | 1.053 | 0.407 | 3.45 |
| *Panel C: Vaccinated/want to get vaccinated* | | | | | | | | | |
| Years of Education | 0.00681*** | -0.0272 | | 0.00646*** | -0.0380 | | 0.0000315 | 0.00189 | |
| | (0.00109) | (0.0410) | | (0.000910) | (0.0343) | | (0.0000308) | (0.00917) | |
| Compulsory Education | | | 0.154*** | | | 0.178*** | | | 0.423* |
| | | | (0.0484) | | | (0.0475) | | | (0.253) |
| Constant | 0.685 | | | 0.531 | | | 0.790*** | | |
| | (0.429) | | | (0.381) | | | (0.247) | | |
| Observations | 9676 | 9676 | 9676 | 13211 | 13211 | 13211 | 30354 | 30354 | 30354 |
| R-squared | 0.249 | -0.147 | | 0.242 | -0.255 | | 0.198 | -0.090 | |
| F-statistic | 13.27 | 0.208 | 10.12 | 17.56 | 0.679 | 14.03 | 0.453 | 0.113 | 2.79 |

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parenthesis. All regressions include country and birth cohort fixed effects. All regressions control for age. Standard errors are clustered at country-by-birth-cohort level..

# 8 Discussion

## 8.1 Relation to previous studies and the theoretical framework

The results in this paper suggest that education decreases smoking. This result is in line with Heckman et al. (2013), Conti et al. (2016), and Etilé and Jones (2011), but the magnitude of the estimate in this paper is larger than the comparable estimate found in Etilé and Jones (2011). Following the reasoning of Brunello et al. (2016), it is possible that the larger estimate in this study stems from the fact that I consider the life history of smoking instead of smoking prevalence. Even though the robustness check indicates that the effect might be overestimated, the result adds to the evidence that more education has a negative impact on smoking.

The large negative effect of education on alcohol is more surprising, given that most studies find no effect of education on alcohol consumption (Clark and Royer, 2013; Braakman, 2011; Park and Kang, 2008). But the result is partly in line with Conti et al. (2016), finding a negative effect for women at the age of 27, but not at the age of 40. The estimated effect on alcohol consumption seems to be large, especially for women, and since the descriptive statistics show that the range of alcoholic units is large it is possible that the magnitude of the effect is driven by outliers. Nevertheless, the results indicate that staying in school longer has a negative effect on the consumption of alcohol.

Even if the results that education decrease smoking and alcohol consumption is not completely in line with previous studies, they are in line with the theoretical framework and give important insights regarding the relative strength of the health cost and the wealth effect. The negative effects of education indicate that the health cost effect dominates the wealth effect regarding the consumption of cigarettes and alcohol. Consumers seem to consider these goods enough harmful to their health to avoid them. Given that the health cost effect takes into account future income, the importance of this effect should decline with age, as the expected years to live decrease. It is thus reasonable that the health cost effect is more prominent when considering the history of smoking than when only considering current smoking prevalence. This could be an additional explanation for the estimated effect of education on smoking being larger in this study than in Etilé and Jones (2011).

Regarding the small negative effect of one additional year of education on the consumption of fruits and vegetables, it is similar to Braakman (2011) and Clark and Royer (2013) but not in line with the studies suggesting that education leads to a more healthy diet (Atella and Kopinska, 2014; Li and Powdthavee, 2015). The theoretical framework also predicts that education would increase the consumption of fruits and vegetables since it would both lead to a positive wealth and health cost effect. The results in this paper give empirical evidence that an increased income due to more education does not seem to encourage increased consumption of fruits and vegetables, contrary to what the theoretical model predicts. Furthermore, the negative health effect of a low intake of fruits and vegetables does not seem to be considered enough harmful to the health to encourage more educated individuals to consume more of these goods through the health cost effect.

The insignificant effects of education on physical activity and the Covid-19 related health behaviors are also in contrast to the theoretical framework. For both physical activity and the Covid-19 related health behaviors, it is unlikely to exist a large wealth effect, since there is no large cost to adopting these behaviors. The net effect should thus be positive for these health behaviors since it should only consist of the health cost effect. One reason for not finding an effect of education on physical activity could be that inactivity is not considered enough harmful to the health. Increasing physical activity, therefore, does not reduce the health cost enough to dominate the opportunity cost of spending time on physical activity. The result that more education does not seem to increase physical activity is in line with the findings in Clark and Royer (2013). Regarding the Covid-19 related health behaviors, taking on these behaviors would likely reduce the risk of getting infected by Covid-19 and thus avoid bad health. Due to the high risk of serious complications of a Covid-19 infection for individuals at a high age, it is not likely that they consider the health cost too low to care about their health behaviors, which the results suggest. Furthermore, there is no evident opportunity cost to these health behaviors. The results thus indicate that there might exist some factors attenuating the effect of education on these health behaviors.

## 8.2 Possible explanations for the variation in the results

A potential reason for the insignificant results for the Covid-19 related health behaviors is that many countries introduced requirements, or strong recommendations, for individuals to wear a facemask in public and get the vaccine. Washing of hands was also a health behavior that was strongly advocated. It is possible that this attenuated the effect of education on these health behaviors. This could also explain the large share of individuals in the Covid-19 sample that reported washing hands more often than before the pandemic, frequently using a facemask, and wanting to get vaccinated against Covid-19.

Another possible reason for the low variation in Covid-19 related health behaviors in the Covid-19 sample, as well as for finding no significant effect of education on these behaviors, is that the differences by education for the main health behaviors are due to an information advantage for more educated individuals. During the Covid-19 pandemic, information about the virus and how to avoid contamination was constantly published.

This can have attenuated differences in health behaviors that would have appeared if this information was less available. The negative health effects of smoking and drinking are however also well known today and an effect of education on these health behaviors is still found.

It is also possible that the differences in significance across the results for the main health behaviors and the Covid-19 related health behaviors are due to differences in the time horizon of the effect on health. It might be that education makes individuals more able to take a longer time perspective into consideration. If that is the case, then education could affect health behaviors for which the effect on health takes place in the long run, such as smoking or alcohol consumption, while there would be no visible effect of education on health behaviors for which the effect on health takes place in the short run, such as getting infected by the Covid-19 virus.

It seems like education can have an effect on reducing health behaviors that directly have a negative impact on health. The results suggest that more education reduces health behaviors that are risky in their nature such as smoking or drinking alcohol. In contrast, education does not seem to have a positive effect on adopting behaviors that would lead to better health. Increasing physical activity, and eating more fruits and vegetables are examples of such behaviors along with washing hands, wearing a facemask, and getting the vaccine during the Covid-19 pandemic.

## 8.3 Heterogeneous effects

A trend that becomes visible in the results for heterogeneous effects is that the instrument seems to be weaker for women than for men. The effect of one year of additional compulsory schooling on the total year of education is lower for women than for men and the F-statistic reported for the excluded instrument in the first stage is generally lower for women than for men. This suggests that this instrument might be better for men than for women. It also indicates that men's schooling decision is more responsive to changes in compulsory schooling than women's and that there, therefore, are more compliers among men than among women.

The results for the effect of education on smoking and alcohol seem to be stronger among women than among men. This finding is in line with Galama et al. (2018) since it confirms the existence of heterogeneous effects, but is in contrast to their conclusion that the effect of education on health behaviors, in general, is weaker for women. It is therefore worth mentioning that the large estimates found in this study could be a result of the instrument being weaker for women than for men. There are however some studies that find larger effects of education on different health behaviors for women than for men.

The evidence of education having a protective effect on BMI, overweight and obesity for women but not for men is interesting in relation to the results in this study (Atella and Kopinska, 2014; Brunello et al. 2013a). The larger negative effect of education on alcohol consumption among women than among men indicates that this could be a potential mechanism explaining the protective effect of education on BMI. The results also suggest that changes in health-seeking behavior or labor force participation due to pregnancy do not seem to play an important role in alcohol consumption or smoking behavior. The results also contradict a normative incentive for smoking and alcohol consumption among successful women in the labor market during recent decades.

# 9 Conclusion

This paper uses an IV approach to study the effect of education on health behaviors. Specifically, I exploit exogenous variation in years of compulsory schooling within and across European countries, following compulsory schooling reforms in the 20th century. The results suggest that among individuals in Europe at the age of 50 or older, one additional year of education decreases smoking and alcohol consumption. I also find a small negative effect of education on the consumption of fruits and vegetables. The results indicate that education does not have an effect on physical activity or Covid-19 related health behaviors.

The differences in significance, magnitude and sign between the OLS estimates and the estimates from the IV regressions accentuate the importance of using an empirical setup that overcomes the endogeneity problem stemming from OVB when estimating the effect of education on health behaviors. Using exogenous variation in compulsory schooling in this setting seems to create a valid instrument for education and the results are in general robust to variations in the bandwidth. The internal validity of the results thus seems to be strong and the results in this study add important insights into the effect of education on health behaviors.

The multi-country setup in this study makes the results relevant for many settings. Using data from several European countries over a long period of time makes the results more general than when exploiting variation within one country and across specific birth cohorts. The external validity of the results in this study can thus be considered stronger than the results of studies using exogenous variation in compulsory education stemming from one specific compulsory schooling reform. When generalizing the results, it is however important to remember that the estimated effect is the effect among the compliers. It is possible that the effect of one more year of schooling is larger for compliers than for individuals that would choose considerably more education than the number of compulsory years of schooling regardless of the reform.

The results are highly policy-relevant even though they estimate the effect among compliers. In line with the theory behind the instrument, these are the individuals that are affected by compulsory schooling reforms. Implementing these health behavior effects in education policy is thus of great importance since it can be used to understand the effect of a compulsory schooling reform on the relevant group of individuals. Furthermore, since the compliers often are low-income individuals, these are the individuals that are usually targeted by a policy aiming to decrease health disparities. The results thus show that education to some extent can be used as a measure to limit disparities in health behaviors and indirectly also in health. Since education does not seem to lead to the adoption of new behaviors in order to improve health, other policies are needed to improve this type of health behavior. Given the results for the Covid-19 related health behaviors, there are some signs that increased information about health effects could help attenuate disparities in health behaviors. This type of information intervention could therefore be used instead of, or in combination with, education policies to decrease health disparities. Since education does not seem to affect the adoption of protective health behaviors in an *at-risk* situation such as the Covid-19 pandemic, other measures need to be taken to limit health disparities in this kind of situation.

The mechanism of the effect of education on health is still unsettled. More research is needed on the effect of education on health behaviors in order to understand through which mechanisms the education-health gradient runs. Additional research is also needed to understand why education has an impact on some health behaviors, but not on others. Knowledge of this would make it possible to tailor policies to increase healthy behaviors and decrease the growing health disparities that are observed around the world.

# 10 References

Amin, V., Behrman, J. R. and Spector, T. D. (2013). Does more schooling improve health outcomes and health related behavior? Evidence from U.K. twins. *Economics of Education Review*, vol. 35, pp. 134-148. Available online: https://doi.org/10.1016/j.econedurev.2013.04.004 [Accessed: 7 May 2022]

Amos, A., Haglund, M. (2000). From social taboo to "torch of freedom": the marketing of cigarettes to women. *Tobacco Control*, vol. 9, no. 1, pp. 3-8. DOI: 10.1136/tc.9.1.3 [Accessed: 2 May 2022]

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

Atella, V. and Kopinska, J. (2014). Body weight, eating patterns, and physical activity: the role of education. *Demography*, vol. 51, no. 4, pp. 1225–1249. Available online: https://doi.org/10.1007/s13524-014-0311-z [Accessed: 4 May 2022]

Braakmann, N. (2011). The causal relationship between education, health and health related behavior: Evidence from a natural experiment in England. *Journal of Health Economics*, vol. 30, no. 4, pp. 753-763. Available online: https://doi.org/10.1016/j.jhealeco.2011.05.015 [Accessed: 7 May 2022]

Brunello, G., Fabbri, D. and Fort, M. (2013a). The causal effect of education on body mass: evidence from Europe. *Journal of Labor Economics*, vol. 31, no. 1, pp. 195-223. Available online: https://doi.org/10.1086/667236 [Accessed: 25 April 2022]

Brunello, G., Fort, M., Weber, G. and Weiss, C.T. (2013b). Testing the Internal Validity of Compulsory School Reforms as Instrument for Years of Schooling, *IZA Discussion Paper Series*, no. 7533. Available online: https://docs.iza.org/dp7533.pdf [Accessed: 28 April 2022]

Brunello, G., Fort, M., Schneeweis, N. and Winter–Ebmer, R. (2016). The causal effect of education on health: what is the role of health behaviors? *Health Economics*, vol. 25, no.3, pp. 314–336. Available online: https://doi.org/10.1002/hec.3141 [Accessed 3 May]

Börsch-Supan, A. (2022a). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1. Release version: 8.0.0. SHARE-ERIC. Data set. DOI:10.6103/SHARE.w1.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022b). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w2.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022c). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w4.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022d). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w5.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022e). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022f). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w7.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022g). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w8.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022h). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. COVID-19 Survey 1. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w8ca.800 [Accessed: 3 April 2022]

Börsch-Supan, A. (2022i). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9. COVID-19 Survey 2. Release version: 8.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w9ca.800 [Accessed: 3 April 2022]

Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *The American Economic Review*, vol. 103, no. 6, pp. 2087-2120. DOI: 10.1257/aer.103.6.2087 [Accessed: 7 May 2022]

Conti, G., Heckman, J. and Pinto, R. (2016). The effects of two influential early childhood interventions on health and healthy behaviour, *The Economic Journal*, vol. 126, no. 596, pp. F28-F65. DOI: 10.1111/ecoj.12420 [Accessed: 14 May 2022]

Cutler, D.M. and Glaeser, E. (2005). What explains differences in smoking, drinking, and other health related behaviors, *American Economic Review*, vol. 95, no. 2, pp. 238-242. DOI: 10.1257/000282805774670464 [Accessed: 13 May 2022].

Cutler, D.M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of Health Economics*, vol 29, no. 1, pp. 1-28.
DOI: 10.1016/j.jhealeco.2009.10.003 [Accessed: 9 May 2022]

Cutler, D.M. and Lleras-Muney, A. (2014). Education and Health: Insights from International Comparisons, in Anthony J. Culyer (ed.), *Encyclopedia of Health Economics*, Vol 1. San Diego: Elsevier, pp. 232-245.

Droomers, M., Schrijvers, C.T.M., Stronks, K., van de Mheen, D. and Mackenbach, J.P. (1999). Educational differences in excessive alcohol consumption: the role of psychosocial and material stressors, *Preventive Medicine*, vol. 29, no. 1, pp. 1-10. DOI: 10.1006/pmed.1999.0496 [Accessed: 10 May 2022]

Etilé F. and Jones, A.M. (2011). Schooling and smoking among the baby boomers – An evaluation of the impact of educational expansion in France, *Journal of Health Economics*, vol. 30, nr. 4, pp. 811-831. Available online: https://doi.org/10.1016/j.jhealeco.2011.05.002 [Accessed: 4 May 2022]

Galama T.J., Lleras-Muney, A. and van Kippersluis, H. (2018). The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence. *NBER Working Paper*, No. 24225, DOI: 10.3386/w24225 [Accessed: 12 May 2022]

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, vol. 80, no. 2, pp. 223-255. Available online: https://www.jstor.org/stable/1830580 [Accessed: 28 April 2022]

Grossman, M. (2000). The Human Capital Model. In: Culyer, A.J. and Newhouse, J.P. (Eds.), *Handbook of Health Economics*, vol. 1, Amsterdam ; New York : Elsevier, 2000, pp. 347-408

Grossman, M. (2006). Education and nonmarket outcomes. In: Hanushek, E. and Welch, F. (Eds.), *Handbook of the Economics of Education*, vol. 2, Amsterdam: North-Holland, pp. 577-633

Hawkins, R.B, Charles, E.J. and Mehaffey, J.H. (2020). Socio-economic status and COVID-19-relatied cases and fatalities, *Public Health*, vol. 189, pp.129-134.
DOI: 10.1016/j.puhe.2020.09.016 [Accessed: 5 May 2022]

Heckman J., Pinto, R. and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, vol. 103, no. 6, pp. 2052-2086. DOI: 10.1257/aer.103.6.2052 [Accessed: 7 May 2022]

Hofmarcher, T. (2021). The effect of education on poverty: A European perspective, *Economics of Education Review*, vol. 83, DOI: 10.1016/j.econedurev.2021.102124 [Accessed: 4 April 2022]

Hummer, R. A., Hernandez, E. M. (2013). The effect of educational attainment on adult mortality in the United States, *Population Bulletin*, vol. 68, no. 1, pp. 1-16. Available online: https://www.prb.org/wp-content/uploads/2013/07/Population-bulletin-2013-68-1-us-education-mortality.pdf [Accessed: 18 May 2022]

Kemptner, D., Jürges, H. and Reinhold, S. (2011). Changes in compulsory schooling and the causal effect of education on health: evidence from Germany, *Journal of Health Economics*, vol. 30, no. 2, pp. 340-354. DOI: 10.1016/j.jhealeco.2011.01.004 [Accessed: 6 May 2022]

Li, J. and Powdthavee, N. (2015). Does more education lead to better health habits? Evidence from the school reforms in Australia, *Social Science Medicine*, vol. 127, pp. 83-91. DOI: 10.1016/j.socscimed.2014.07.021 [Accessed: 5 May 2022]

Lundborg, P. (2013). The health returns to schooling - what can we learn from twins?, *Journal of Population Economics*, vol. 26, pp. 673-701. Available online: https://doi.org/10.1007/s00148-012-0429-5 [Accessed: 7 May 2022]

Lundborg, P., Plug, E. and Wurtz Rasmussen, A. (2017). Can women have children and a career? IV Evidence from IVF treatments. *American Economic Review*, vol. 107, no. 6, pp. 1611–1637. Available online: https://doi.org/10.1257/aer.20141467 [Accessed: 9 May 2022]

Meara, E.R., Richards, S., Cutler, D.M. (2008). The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. *Health Affairs*, vol. 27, no. 2, pp. 350–360. Available online: https://doi.org/10.1377/hlthaff.27.2.350 [Accessed: 18 May 2022]

Mokdad, A. H., Marks, J. S., Stroup, D. F. and Gerberding, J.L. (2004). Actual causes of death in the United States, *Journal of the American Medical Association*, vol. 291, nr. 10, pp. 1238-1245. DOI: 10.1001/jama.291.10.1238 [Accessed: 19 May 2022]

OECD/Eurostat/UNESCO Institute for Statistics (2015). ISCED 2011 Operational Manual: Guidelines for Classifying National Education Programmes and Related Qualifications, Paris: OECD Publishing. Available online: https://www.oecd-ilibrary.org/docserver/9789264228368-en.pdf?expires=1653325839&id=id&accname=ocid177253&checksum\protect\@normalcr\relax=0663BDA43252B57459AC0B57B44B6E07 [Accessed: 4 May 2022]

Park, C. and Kang, C. (2008). Does education induce healthy lifestyle?, *Journal of Health Economics*, vol. 27, no. 6, pp. 1516-1531. DOI: 10.1016/j.jhealeco.2008.07.005 [Accessed: 3 May 2022]

SHARE (2022a), SHARE - The Survey of Health, Ageing and Retirement in Europe. Available online: http://www.share-project.org/home0.html [Accessed: 20 April 2022]

SHARE (2022b). Release Guide 8.0.0. Available online: http://www.share-project.org/fileadmin/pdf_documentation/SHARE_release_guide_8-0-0.pdf [Accessed: 11 April 2022]

Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments, *Econometrica*, vol. 65, no. 3, pp. 557-586. Available online: https://doi.org/10.2307/2171753 [Accessed: 10 May 2022]

Van Kippersluis, H. and Galama, T. J. (2014). Wealth and health behavior: Testing the concept of a health cost, *European Economic Review*, vol. 72, pp. 197-220. DOI: 10.1016/j.euroecorev.2014.10.003 [Accessed: 11 May 2022].

# A Appendix

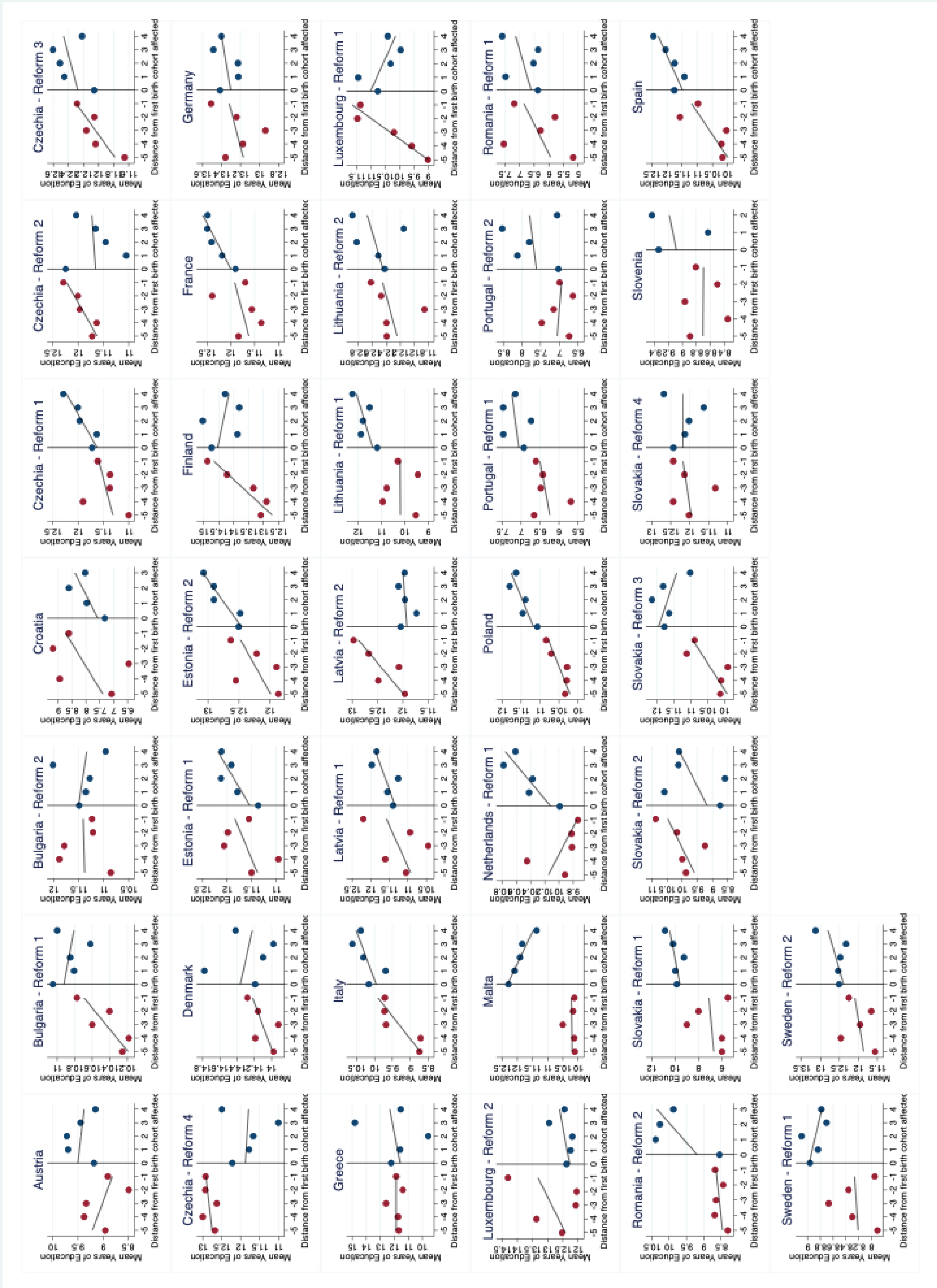
Table A.1: Compulsory schooling reforms

| Country | Years of compulsory schooling | School starting age | School leaving age | Reform date | First cohort affected |
|---|---|---|---|---|---|
| Austria | 8 to 9 | 6 | 14 to 15 | 1962 | 1952,00 |
| Bulgaria 1 | 7 to 8 | 7 | 14 to 15 | 1959 | 1946,00 |
| Bulgaria 2 | 8 to 9 | 7 | 15 to 16 | 1969 | 1958,00 |
| Croatia 1 | 4 to 7 | 7 | 11 to 14 | 1945 | 1935,00 |
| Czechia 1 | 8 to 9 | 6 | 14 to 15 | 1948 | 1934,00 |
| Czechia 2 | 9 to 8 | 6 | 15 to 14 | 1953 | 1939,00 |
| Czechia 3 | 8 to 9 | 6 | 14 to 15 | 1960 | 1947,00 |
| Czechia 4 | 9 to 10 | 6 | 15 to 16 | 1978 | 1964,00 |
| Denmark 1 | 7 to 9 | 7 | 14 to 16 | 1972 | 1958,00 |
| Estonia 1 | 7 to 8 | 7 | 14 to 15 | 1958 | 1945,00 |
| Estonia 2 | 8 to 11 | 7 | 15 to 18 | 1970 | 1955,00 |
| Finland | 6 to 9 | 7 | 13 to 16 | 1968 | 1963,00 |
| France | 8 to 10 | 6 | 14 to 16 | 1959 | 1953,00 |
| Germany | 8 to 9 | 6 | 14 to 15 | 1967 | 1953,00 |
| Greece | 6 to 9 | 6 | 12 to 15 | 1976 | 1964,00 |
| Italy 1 | 5 to 8 | 6 | 11 to 14 | 1962 | 1952,00 |
| Latvia 1 | 7 to 8 | 7 | 14 to 15 | 1958 | 1945,00 |
| Latvia 2 | 8 to 11 | 7 | 15 to 18 | 1970 | 1955,00 |
| Lithuania 1 | 7 to 8 | 7 | 14 to 15 | 1958 | 1945,00 |
| Lithuania 2 | 8 to 11 | 7 | 15 to 18 | 1970 | 1955,00 |
| Luxembourg 1 | 7 to 8 | 6 | 13 to 14 | 1945 | 1932,00 |
| Luxembourg 2 | 8 to 9 | 6 | 14 to 15 | 1963 | 1950,00 |
| Malta 1 | 8 to 10 | 6 | 14 to 16 | 1974 | 1960,00 |
| Netherlands 1 | 7 to 8 | 6 | 13 to 14 | 1947 | 1937,00 |
| Poland 1 | 7 to 8 | 7 | 14 to 15 | 1961 | 1952,00 |
| Portugal 1 | 3 to 4 | 7 | 12 | 1956/1960* | 1950,00 |
| Portugal 2 | 4 to 6 | 7 | 12 to 14 | 1964 | 1957,00 |
| Romania 1 | 7 to 4 | 7 | 11 | 1948 | 1935,00 |
| Romania 2 | 4 to 7 | 7 | 14 | 1958 | 1947,00 |
| Slovakia 1 | 8 to 9 | 6 | 14 to 15 | 1948 | 1934,00 |
| Slovakia 2 | 9 to 8 | 6 | 15 to 14 | 1953 | 1939,00 |
| Slovakia 3 | 8 to 9 | 6 | 14 to 15 | 1960 | 1947,00 |
| Slovakia 4 | 9 to 10 | 6 | 15 to 16 | 1978 | 1964,00 |
| Slovenia 1 | 4 to 7 | 7 | 11 to 14 | 1945 | 1935,00 |
| Spain 1 | 7 to 8 | 6 | 13 to 14 | 1970 | 1958,00 |
| Sweden 1 | 6 to 7 | 7 | 13 to 14 | 1936 | 1923,00 |
| Sweden 2 | 7 to 9 | 7 | 14 to 16 | 1962 | 1951,00 |

Note: The information on the compulsory schooling reforms is collected from the database of reforms in Hofmarcher (2021).

*The reform was enacted for boys in 1956 and for girls in 1960.

Figure A.1: First stage by reform




Note: Mean years of education by distance from first birth cohort affected for each reform. The sample window is five years before and after each reform.

Table A.2: Selection of reforms

| Country | Coefficient | SE | t-stat | p-value |
|---|---|---|---|---|
| Austria | 0.607 | 0.437 | 1.39 | 0.165 |
| Bulgaria 1 | 0.316 | 0.485 | 0.65 | 0.515 |
| Croatia | -0.861 | 0.389 | -2.21 | 0.028 |
| Czechia 1 | 0.140 | 0.332 | 0.42 | 0.673 |
| Denmark | 0.046 | 0.152 | 0.30 | 0.761 |
| Estonia 1 | -0.450 | 0.331 | -1.36 | 0.174 |
| Finland | -0.224 | 0.184 | -1.22 | 0.224 |
| France | 0.034 | 0.147 | 0.23 | 0.820 |
| Germany | 0.032 | 0.255 | 0.12 | 0.901 |
| Greece | 0.108 | 0.158 | 0.69 | 0.493 |
| Italy | -0.052 | 0.121 | -0.43 | 0.666 |
| Latvia 1 | -0.541 | 0.742 | -0.73 | 0.466 |
| Lithuania 1 | 1.043 | 0.610 | 1.71 | 0.088 |
| Luxembourg 1 | -0.102 | 1.086 | -0.09 | 0.926 |
| Malta | 1.087 | 0.282 | 3.86 | 0.000 |
| Netherlands 1 | 0.541 | 0.353 | 1.53 | 0.125 |
| Poland | 0.340 | 0.243 | 1.40 | 0.162 |
| Portugal 1 | 0.466 | 0.629 | 0.74 | 0.459 |
| Romania 1 | 0.105 | 0.331 | 0.32 | 0.752 |
| Slovakia 1 | 2.210 | 1.327 | 1.66 | 0.101 |
| Slovenia | 0.115 | 0.166 | 0.70 | 0.487 |
| Spain | -0.098 | 0.408 | -0.24 | 0.811 |
| Sweden 1 | 0.568 | 0.576 | 0.99 | 0.324 |
| Bulgaria 2 | 0.368 | 0.490 | 0.75 | 0.453 |
| Czechia 2 | 0.602 | 0.297 | 2.03 | 0.043 |
| Estonia 2 | -0.062 | 0.091 | -0.68 | 0.499 |
| Latvia 2 | -0.352 | 0.188 | -1.87 | 0.063 |
| Lithuania 2 | -0.020 | 0.158 | -0.13 | 0.900 |
| Luxembourg 2 | -1.315 | 0.717 | -1.83 | 0.067 |
| Portugal 2 | 0.202 | 0.306 | 0.66 | 0.509 |
| Romania 2 | 0.133 | 0.222 | 0.60 | 0.551 |
| Slovakia 2 | 0.004 | 0.616 | 0.01 | 0.994 |
| Sweden 2 | 0.058 | 0.152 | 0.38 | 0.703 |
| Czechia 3 | -0.106 | 0.120 | -0.53 | 0.596 |
| Slovakia 3 | 0.384 | 0.516 | 0.74 | 0.460 |
| Czechia 4 | -0.922 | 0.684 | -1.35 | 0.178 |
| Slovakia 4 | 0.109 | 0.329 | 0.33 | 0.740 |

Note: The sample window is five years before and after each reform.

Table A.3: Descriptive statistics by gender - Main sample

| | Men | | | | | Women | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs | Mean | Std. Dev. | Min | Max | Obs | Mean | Std. Dev. | Min | Max |
| *Education* | | | | | | | | | | |
| Years of compulsory education | 14,332 | 7.722 | 1.468 | 3 | 11 | 17,445 | 7.747 | 1.482 | 3 | 11 |
| Years of education | 14,332 | 11.556 | 4.030 | 0 | 25 | 17,445 | 10.939 | 3.906 | 0 | 25 |
| *Socioeconomic characteristics* | | | | | | | | | | |
| Age | 14,332 | 65.530 | 8.362 | 50 | 99 | 17,445 | 65.482 | 8.640 | 50 | 99 |
| Male | 14,332 | 1 | 0 | 1 | 1 | 17,445 | 0 | 0 | 0 | 0 |
| Year of birth | 14,332 | 1949.231 | 8.192 | 1918 | 1968 | 17,445 | 1949.543 | 8.337 | 1918 | 1968 |
| Born in the country of interview | 14,332 | 0.992 | 0.086 | 0 | 1 | 17,445 | 0.992 | 0.088 | 0 | 1 |
| In good subjective health | 14,031 | 0.630 | 0.483 | 0 | 1 | 17,123 | 0.614 | 0.487 | 0 | 1 |
| *Health behaviors* | | | | | | | | | | |
| Units of alcoholic beverage last 7 days | 8,653 | 6.35583 | 9.410 | 0 | 140 | 11,121 | 2.057 | 4.579 | 0 | 120 |
| Ever smoked daily | 8,301 | .571136 | 0.495 | 0 | 1 | 10,281 | 0.356 | 0.479 | 0 | 1 |
| Vigorous physical activity (at least once/week) | 12,935 | .4964824 | 0.500 | 0 | 1 | 15,851 | 0.434 | 0.496 | 0 | 1 |
| Eat fruits/vegetables (at least three times/week) | 8,279 | .8761928 | 0.329 | 0 | 1 | 10,219 | 0.926 | 0.262 | 0 | 1 |

Note: The samples include up to five birth cohorts before and after a reform.

Table A.4: Descriptive statistics by gender - Covid-19 sample

| | Men | | | | | Women | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Obs | Mean | Std. Dev. | Min | Max | Obs | Mean | Std. Dev. | Min | Max |
| *Education* | | | | | | | | | | |
| Years of compulsory education | 5,841 | 7.627 | 1.618 | 3 | 11 | 7,885 | 7.657 | 1.632 | 3 | 11 |
| Years of education | 5,841 | 11.649 | 4.137 | 0 | 25 | 7,885 | 10.974 | 3.977 | 0 | 25 |
| *Socioeconomic characteristics* | | | | | | | | | | |
| Age | 5,841 | 66.442 | 7.120 | 50 | 95 | 7,885 | 66.462 | 7.665 | 50 | 99 |
| Male | 5,841 | 1 | 0 | 1 | 1 | 7,885 | 0 | 0 | 0 | 0 |
| Year of birth | 5,841 | 1951.121 | 6.684 | 1924 | 1967 | 7,885 | 1951.116 | 7.251 | 1921 | 1968 |
| Born in the country of interview | 5,841 | 0.993 | 0.084 | 0 | 1 | 7,885 | 0.994 | 0.079 | 0 | 1 |
| In good subjective health | 5,835 | 0.661 | 0.473 | 0 | 1 | 7,877 | 0.634 | 0.482 | 0 | 1 |
| *Health behaviors* | | | | | | | | | | |
| Units of alcoholic beverage last 7 days | 5,256 | 6.276 | 9.429 | 0 | 130 | 7,139 | 1.978 | 4.542 | 0 | 120 |
| Ever smoked daily | 4,187 | 0.533 | 0.499 | 0 | 1 | 5,675 | 0.331 | 0.471 | 0 | 1 |
| Vigorous physical activity (at least once/week) | 5,471 | 0.511 | 0.500 | 0 | 1 | 7,425 | 0.448 | 0.497 | 0 | 1 |
| Eat fruits/vegetables (at least three times/week) | 4,303 | 0.880 | 0.325 | 0 | 1 | 5,838 | 0.924 | 0.266 | 0 | 1 |
| *Covid-19 related health behaviors* | | | | | | | | | | |
| Wear facemask often | 5,052 | 0.813 | 0.390 | 0 | 0 | 6,506 | 0.842 | 0.365 | 0 | 1 |
| Wash hands more often than before the pandemic | 5,710 | 0.879 | 0.326 | 0 | 1 | 7,784 | 0.897 | 0.304 | 0 | 1 |
| Want to get vaccinated against covid-19 | 4,806 | 0.858 | 0.349 | 0 | 1 | 6,695 | 0.839 | 0.368 | 0 | 1 |

Note: The samples include up to five birth cohorts before and after a reform.