Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

hche | Hamburg Center for Health Economics

Unlucky to Be Young? The Long-Term Effects of School Starting Age on Smoking Behaviour and Health

Michael Bahrs Mathias Schumann

Research Paper
Year: 2016
No: 13

Unlucky to Be Young? The Long-Term Effects of School Starting Age on Smoking Behaviour and Health

Michael Bahrs* Mathias Schumann†

August 4, 2016

Abstract: In this study, we analyse the long-term effects of school starting age on smoking behaviour and health in adulthood. School entry rules combined with birth month are used as an instrument for school starting age. The analysis adopts the German Socio-Economic Panel data and employs a fuzzy regression discontinuity design. The results reveal that school starting age reduces the long-term risk to smoke, improves long-term health, and affects physical rather than mental health. In addition, we find that the relative age composition of peers and the school environment are important mechanisms.

Keywords: smoking, health, peer effects, education, school starting age, regression discontinuity design.

JEL Classification: I12, I21, I28.

Acknowledgments: We thank Sarah Dahmann, Dan Dammann, Peter Eibich, Jan Marcus, Annemarie Paul, Martin Salm, Thomas Siedler, Felix Weinhardt, and the participants of the 12th International German Socio-Economic Panel User Conference and the 11th European Conference on Health Economics for their helpful comments. This work was supported by the Federal Ministry of Education and Research (*Bundesministerium für Bildung und Forschung*) through the project '*Nicht-monetäre Erträge von Bildung in den Bereichen Gesundheit, nicht-kognitive Fähigkeiten sowie gesellschaftliche und politische Partizipation*'. Opinions expressed in this paper are those of the authors. Any remaining errors are our own.

*Universität Hamburg, Fakultät für Wirtschafts- und Sozialwissenschaften, Von-Melle-Park 5, 20146 Hamburg, Germany, Email: michael.bahrs@wiso.uni-hamburg.de.

†Universität Hamburg, Fakultät für Wirtschafts- und Sozialwissenschaften, Von-Melle-Park 5, 20146 Hamburg, Germany, Email: mathias.schumann@wiso.uni-hamburg.de.

1 Introduction

Smoking results in significant healthcare costs (Wacker et al., 2013; Xu et al., 2015) and is considered the leading cause of preventable deaths (WHO, 2015a). According to the World Health Organization, in 2013, the share of tobacco smokers among persons aged 15 years and above was 18.1% in the US, 20.3% in the UK and 30.7% in Germany (WHO, 2015b).[1] It is, thus, imperative to understand the determinants of smoking behaviour, particularly for policymakers, to reduce the prevalence of smoking, and thereby, improve the health status of the population and decrease smoking-related healthcare costs.

Smoking habits are generally formed during childhood and adolescence and persist into adulthood. In Germany, the average age to start smoking was 17.3 years among 35–39-year-olds in 2013 (Destatis, 2014). In the US, 88.2% of adults who had smoked daily at some point reported trying their first cigarette by the age of 18 years (U.S. Department of Health and Human Services, 2012). Gruber and Zinman (2000) show that adolescent smoking is a strong predictor of adult smoking and a percentage point increase in adolescent smoking translates into a 0.25–0.5 percentage points higher likelihood to smoke by those adolescents as adults. Chassin et al. (1996) find that smoking rates do not significantly decline among those in their late twenties and this pattern is stable across birth cohorts. Thus, analysing factors determining smoking behaviour in early life is important to prevent adolescents from smoking in their adulthood.

Sacerdote (2011) compiles a literature review and points out the role of school peers in social outcomes such as smoking and health. Norton et al. (1998), Gaviria and Raphael (2001) and Powell et al. (2005) find that an increase in the share of student smokers in school increases an individual's risk to smoke in adolescence. A recent strand of literature analyses the effects of individual school starting age on social outcomes in adolescence.[2] Students who start school relatively young are exposed to the behaviour of older class peers. School starting age, therefore, does affect social outcomes through relative age differences among class peers. Related studies have examined the effects of school starting age on several outcomes including non-cognitive skills, educational attainment and labour market outcomes (e.g. Bedard and Dhuey, 2006; Puhani and Weber, 2007; McEwan and Shapiro, 2008; Elder and Lubotsky, 2009; Dobkin and Ferreira, 2010; Mühlenweg and Puhani, 2010; Black et al., 2011; Mühlenweg et al., 2012; Fredriksson and Öckert, 2014; Dustmann et al., 2016; Landersø et al., 2016).

However, few studies have analysed the impact of school starting age on smoking behaviour and health. Argys and Rees (2008) find that female adolescents who enrol in

[1]The World Health Organization standardises national smoker rates by applying age-specific smoker rates by sex in each population to a statistical standard population to enable cross-country comparisons.

[2]This branch of literature and our study exploit legal school starting age cut-offs to analyse the effects of relative differences in individual school starting age. By contrast, Fletcher and Kim (2016) analyse the effects of shifts in school entry cut-offs that change the general school starting age.

school at a relatively young age face a higher risk of smoking in grades 6–12. Black et al. (2011) find that 18-year-old male conscripts who started school relatively young show slightly poorer mental health in military medical inspection. Several studies show that young school starters are more likely to be diagnosed with attention-deficit/hyperactivity disorder (ADHD) in childhood and adolescence (Elder and Lubotsky, 2009; Elder, 2010; Evans et al., 2010; Morrow et al., 2012; Schwandt and Wuppermann, 2016).[3] Exploiting school entry cut-off dates, Anderson et al. (2011) show that an additional year of education does not impact children's body mass index or their likelihood of being obese.

Despite the contributions of these studies, the evidence on school starting age effects on smoking behaviour and health remains relatively sparse. Moreover, the literature focuses on the short-term effects of school starting age in adolescence and young adulthood owing to data restrictions. Thus, whether the effects of school starting age on smoking behaviour and health in adolescence persist into adulthood remains an open question. From a policy perspective, it is important to determine whether the effects of school starting age remain stable or vanish over time.

In this study, we go beyond adolescence and examine the causal long-term effects of school starting age on smoking and health among adults in their late thirties. We employ a fuzzy regression discontinuity design to account for the endogeneity of school starting age because some parents time their children's school enrolment with respect to (unobserved) child characteristics, such as preschool health and perceived school readiness. Exogenous cut-off dates for school entry, as per which a child must be six years old to enter primary school, are used as an instrument for school starting age. The analysis utilises survey data from the German Socio-Economic Panel (SOEP). We find that an increase in school starting age by one month reduces the long-term risk of smoking by about 1.3 percentage points (4%) and increases the long-term likelihood of reporting good or very good health by about 1.6 percentage points (2.4%). The effects on self-rated health can be explained on the basis of changes in physical rather than mental health. Moreover, an increase in school starting age lowers the average age of friends in adulthood; we interpret this result as evidence that suggests the importance of school peers' age composition as a mechanism through which school starting age affects smoking behaviour and health. Furthermore, we show that school environment partly explains the effects of school starting age by exploiting the association between the type of secondary school degree and peer smoking intensity.

Our study makes the following three contributions to the literature. First, we complement the literature on the effects of school starting age on smoking behaviour, overall health, and physical and mental health. Second, we are able to estimate long-term effects

[3]These studies interpret the higher number of diagnoses among younger school starters as misdiagnoses, which is confirmed by Dalsgaard et al. (2012).

of school starting age on these outcomes. Third, we shed light on mechanisms through which school starting age affects smoking behaviour and health.

School starting age is expected to affect smoking behaviour and health through, first, the age composition of school peers, and second, school environment. School entry cut-off dates create exogenous variation in the relative age composition of class peers. Students born just before the cut-off are supposed to start school one year earlier than those born immediately after the cut-off—thus, students' age in the same class can differ by almost a year. Age is also an important factor affecting smoking behaviour in adolescence. Figure 1 illustrates that the share of smokers in Germany significantly increases by age.[4] The relationship between age and smoking prevalence implies that students who started school relatively young are confronted by peers who smoke earlier. Young school starters are, therefore, more likely to start smoking than old school starters in the long term, because the former are exposed much earlier in life to peers who smoke.

As for school environments, Germany has a school tracking system that assigns students to different secondary school types at the age of 10. Jürges and Schneider (2007), Puhani and Weber (2007), Mühlenweg and Puhani (2010) and Dustmann et al. (2016) show that entering primary school older increases the likelihood of entering secondary schools of higher tracks. High-track schools offer a better school environment than low-track ones. Figure 2 shows that the share of smokers is 20% in low-track schools and 8% in high-track schools among students who are 12–15 years old. Students who started primary school relatively young are thus, more likely to be exposed to a school environment with a higher number of smokers. School starting age is, therefore, expected to increase the risk of smoking and decrease long-term health through both the age composition of class peers and school environments.

The remainder of this paper is structured as follows. Section 2 describes the SOEP data and outcome variables. Section 3 explains the identification strategy, the German school entry rule used as an instrument for school starting age and the validity of the identifying assumptions for the fuzzy regression discontinuity design. Section 4 shows the descriptive statistics. Section 5 presents the main results and several robustness checks. Section 6 discusses potential mechanisms through which school starting age may affect smoking behaviour and health with focus on the role of school peers. Section 7 concludes the paper.

[4]Figures 1 and 2 are based on data from the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), 2003–2006, administered by the Robert Koch Institute. KiGGS is a nationwide clustered random sample of 17,641 children and adolescents (0–17 years) and their parents (Hölling et al., 2012).

2 Data

We use data from the German Socio-Economic Panel (SOEP), which is an annually conducted, representative household survey (Wagner et al., 2007). The SOEP includes about 30,000 individuals living in roughly 11,000 households in Germany. Adult members of the household are interviewed about their socioeconomic and demographic characteristics. The SOEP offers a rich set of information such as income, employment, education and health and has been used to analyse health-related questions such as the effects of public smoking bans on smoking behaviour (Anger et al., 2011), spousal job loss on mental health (Marcus, 2013) and retirement on health (Eibich, 2015).

2.1 Sample

We include respondents who provided complete information regarding the analysis' outcomes and covariates. The sample comprises respondents who grew up in households that participated in the SOEP when they were children or adolescents (about 17.5% of the sample) and respondents who entered the panel after maturity (about 82.5%). We conduct robustness checks to show that our results are robust to the exclusion of adolescents and young adults.

In the main specifications, we use the first available observation of each respondent separately for each outcome and neglect repeated observations for two reasons. First, Eibich (2015) and Godard (2016) show that retirement reduces the likelihood of smoking and improves health. The inclusion of observations close to retirement may, therefore, bias the effect of school starting age on smoking behaviour and health. Second, the use of observations that are closest to a respondent's schooling period allows us to more accurately gauge the mechanisms through which school starting age affects smoking behaviour and health. Thus, we use cross-sectional data comprising respondents interviewed in different survey years at varying ages. The robustness checks show that the main results are robust to both the inclusion of all available observations for each respondent and exclusion of respondents at least 60 years old.

While the literature on the effects of school starting age on smoking behaviour and health is limited to short-term effects owing to data restrictions, the SOEP allows us to look beyond adolescence. The respondents in the estimation sample are, on average, 37 years old.[5] It is, therefore, possible for us to analyse the long-term effects of school starting age on smoking behaviour and health.

[5]Table A-1 shows that respondents' average age at the time of the SOEP interview is about 35.5–37.9 years, depending on the analysed outcome and specification. Furthermore, it shows that respondents' age does not statistically differ between persons born before or after the cut-off.

2.2 Outcomes

We use adult smoking behaviour and subjective and quasi-objective[6] health measures as outcomes to analyse the long-term relationship between smoking behaviour and health and the effect of school starting age on both. Information on health and health-related behaviour is available in the SOEP either annually or for certain waves.[7] Whereas annual self-rated health data are available for 1992–2013, those on smoking behaviour are available from 1998 onwards for roughly every second year. The SF12 measures of physical and mental health are available since 2002 and for every second year. Consequently, the sample size varies from 1,674 to 3,856 in the preferred specifications across outcomes.

To analyse smoking behaviour, we use an indicator variable that takes the value one if the respondent was a smoker at the time of the study and zero otherwise. We adopt the self-rated health, physical health and mental health scores as health measures. For self-rated health, respondents are asked to assess their current health on a five-point scale, where 1 is 'bad' and 5 is 'very good'. Because self-rated health is an ordinal variable, we use an indicator based on self-rated health as an additional outcome for a more intuitive interpretation of the effect. The indicator takes the value of one for 'good' and 'very good' health and zero otherwise.

The physical and mental health scores are taken from the continuous quasi-objective SF12. The SF12 is a concise instrument to measure physical and mental health and is based on a set of 12 questions about various health aspects, including body pain and emotional functioning. The 12 questions are aggregated to eight subscales, which in turn, are used to calculate the physical and mental health scores using an exploratory factor analysis. Both scores are continuous and normalised to have values ranging from 0 to 100, with a mean of 50 and a standard deviation of 10 in the 2004 SOEP sample (Andersen et al., 2007). A higher value indicates better physical or mental health. Studies have found that the SF12 and therefore, the physical and mental health scores are valid and reliable and perform well compared to other brief health measures (Ware et al., 1996; Salyers et al., 2000).

To shed light on potential mechanisms through which school entry age might affect smoking behaviour and health, we analyse three measures on the respondents' social networks as additional outcomes: 1) number of friends 2) average age of friends and 3) relative age of friends (average age of friends divided by respondent's age in years). Both age measures exclude family members and relatives. All three measures are related to the respondent's network at the time of the SOEP interview. The descriptive statistics of the outcomes are shown in Table 1 and detailed in Section 4.

[6] 'Quasi-objective' means that the respective health measure enables health comparisons across different groups of persons (e.g. age groups).

[7] Table A-2 in the Appendix provides an overview of the availability of outcome measures across SOEP waves.

3 Research Design

Our main variable of interest in the analysis is school starting age measured in months. An OLS regression of smoking behaviour and health on school starting age is unlikely to uncover the causal effects of school starting age and would result in biased estimates because school starting age is likely to be endogenous. Parents might determine the school entry age of their children strategically by accounting for factors that are unobserved in the data. They could be concerned about their children's preschool health or health-related factors such as school readiness and thus, might move up or postpone their children's school enrolment (Graue and DiPerna, 2000; Stipek, 2002).

To resolve the endogeneity of school starting age, the economic literature utilises exogenous school entry rules as an instrument. The German school entry rule has been used to study the effects of school starting age on the likelihood of attending higher track schools (Jürges and Schneider, 2007; Mühlenweg and Puhani, 2010), test scores at the end of primary school (Puhani and Weber, 2007), and long-term labour market outcomes (Dustmann et al., 2016).

We adopt a fuzzy regression discontinuity design to analyse the long-term effects of school starting age on smoking behaviour and health outcomes using the German school entry rule as an instrument for school starting age. The school entry rule determines whether a child is supposed to start school in year t or year $t+1$, depending on the month of birth. We estimate a local average treatment effect (LATE), which is the causal effect for compliers, that is, children who start school according to the legal age-based school entry rule (Hahn et al., 2001).

3.1 School Entry Rule and Instrument

In Germany, a child's date of birth determines the intended date of entry into primary school. The academic school year in each federal state of Germany is from August 1st to July 31st.[8] Children who turn six by June 30th in year t are supposed to start primary school on August 1st in year t, while those who turn six on July 1st or after in year t must start primary school on August 1st in year $t+1$.

Note that different school entry cut-offs existed before harmonisation in Germany.[9]

[8]Before the German reunification in 1990, the starting month differed by federal state. Before 1964, the starting month in the Federal Republic of Germany was April or August. However, in 1964, the *Hamburger Abkommen* harmonised the start of primary school to August 1st. The starting month in the German Democratic Republic was September 1st but in 1990, it was also changed to August 1st.

[9]In the Federal Republic of Germany, some federal states had school entry cut-offs other than June 30th (about 21% of the sample), although this was later harmonised with the ratification of the *Hamburger Abkommen* on October 28, 1964. Before the German reunification in 1990, the school entry cut-off in the German Democratic Republic was May 31st (about 21% of the sample). However, following the reunification, the federal states of the former GDR adopted June 30th as the cut-off, which is relevant for about 58% of the sample.

Because older and younger cohorts are pooled in the analysis, the instrument is coded such that different cut-offs are incorporated. The differences in entry cut-offs for older cohorts create additional variation in school starting age. The inclusion of cut-offs other than the June 30th increases the generalisability of our results, because we can rule out seasonal idiosyncrasy.

Parents, however, may still decide to enrol their children later or earlier than the school entry rule stipulates. Nonetheless, there is considerable discontinuity in school starting age at the school entry cut-off, as shown in the upper left graph in Figure 3. The abscissa shows the distance between a person's birth month and school entry cut-off month and the ordinate shows the (observed) average school starting age (in months). Compliance with the school entry rule is not perfect because the jump at the cut-off is less than eleven months. Nonetheless, the school entry rule is a strong instrument as indicated by the considerable discontinuity of about three months at the cut-off and the negative trends to the left and right of the cut-off. Two months after the cut-off, this discontinuity even increases to five months.

Therefore, we use the school entry cut-off to define the binary instrument $older_i$. The instrument takes the value one if the respondent turned six after the cut-off in year t and should have been enrolled in year $t + 1$; it takes the value zero if the respondent turned six before the cut-off in year t and should have been enrolled in year t.

For respondents whose household participated in the SOEP during their childhood or adolescence, direct information for year and federal state of school start and school starting age sa_i is available; by contrast, the same data are unavailable for those who participated in the SOEP after maturity. For the former group, we construct the instrument $older_i$ by combining information on respondents' date of birth, year and federal state of school start. For the latter group, the highest school degree attained, the year in which it was completed, and the federal state where it was completed are used to determine the year and federal state of school start. In combination with respondents' date of birth, we construct the variable school starting age sa_i and instrument $older_i$. We discuss the possibility of measurement error in sa_i and $older_i$ for individuals without direct information regarding school starting age in Section 5 and show that the potential measurement error is negligible.

3.2 Fuzzy Regression Discontinuity Design

We employ a fuzzy regression discontinuity design because compliance with the date of birth cut-off is not perfect. However, we can still use the substantial discontinuity in school starting age at the cut-off as an instrument (Imbens and Lemieux, 2008). The fuzzy regression discontinuity design can be implemented using two-stage least squares estimation (Angrist and Pischke, 2009). In the first stage, observed school starting age

sa_i (measured in months) is regressed on the instrument $older_i$, where subscript i denotes individual i:

$$sa_i = \alpha_0 + \alpha_1 older_i + \pi X_i' + \gamma_b + \gamma_s + \gamma_w + e_i. \quad (1)$$

The estimate for α_1 is expected to be positive. X_i is a vector of covariates predetermined with respect to birth, including respondents' gender, paternal and maternal school education, and migration background.[10] Further, the regression includes birth year indicators, γ_b, indicators for the federal state where the child enrolled in primary school, γ_s, and survey wave indicators γ_w; e_i is an idiosyncratic error term.[11]

In the second stage, the outcome of interest y_i is regressed on predicted school starting age in months $s\hat{a}_i$:

$$y_i = \beta_0 + \beta_1 s\hat{a}_i + \delta X_i' + \gamma_b + \gamma_s + \gamma_w + u_i, \quad (2)$$

where X_i', γ_b, γ_s and γ_w are the same as those in the first stage and u_i is an idiosyncratic error term.

There are two main approaches to implementing a (fuzzy) regression discontinuity design. One can restrict the sample to a narrow bandwidth around the cut-off or use the entire sample and model polynomial trends of the running variable around the cut-off (Lee and Lemieux, 2010). In our main specifications, we restrict the sample to a two-month window around the birth month cut-off (i.e. respondents born one month before and after the cut-off) and a four-month window around the birth month cut-off (i.e. respondents born two months before and after the cut-off). We implement this approach instead of modelling trends for the entire sample because our running variable is discrete rather than continuous, which renders the estimation of flexible trends using polynomials infeasible (Angrist and Pischke, 2009). The advantage of estimating the effects in a narrow window is a reduction in bias because observations close to the cut-off are more comparable in terms of observable and unobservable characteristics. The disadvantage is the loss of precision due to the loss of observations. In the robustness checks, we use the entire sample and include linear trends of the running variable.

3.3 Identifying Assumptions

The two-stage least squares estimate for β_1 uncovers the causal effect of school starting age on the outcome of interest if three assumptions are fulfilled. First is the *relevance*

[10]The highest secondary school degree for the respondents' mothers and fathers are measured using three indicators: 1) upper secondary school degree (*Abitur*) 2) intermediate secondary school degree (*Realschulabschluss*, *Fachoberschulabschluss*) and 3) lower secondary school degree (*Hauptschulabschluss*) or no secondary school degree.

[11]Respondents who started school in the former GDR are assigned a GDR indicator.

assumption: the instrument must be sufficiently partially correlated with school starting age. The first-stage F-statistic is well above the conventional thresholds in each specification in our analysis.[12] Depending on the outcome and specification, the F-statistic is greater than 60 in the two-month window and larger than 200 in the four-month window in our main specifications (see Table 3).

Second is the *exclusion restriction*: birth month has no direct effect on smoking behaviour and health. The instrument should affect the outcomes only through school starting age. In contrast to the US, there is no interaction between school entry age and compulsory school leaving laws in Germany. Students in the US may leave school on their 16th birthday, and thus, the date of birth affects the length of formal schooling. In contrast, students in Germany must complete nine years of schooling, irrespective of their date of birth and therefore, the length of formal schooling is unaffected.

Third is the *independence assumption*: respondents' date of birth is random around the school entry cut-off. Randomness implies that parents do not systematically manipulate their children's date of birth with respect to the school entry cut-off. The advantage of using birth month as an instrument rather than birth quarter is that strategic birth timing is more unlikely between adjacent months than between adjacent seasons.[13]

Figure 4 suggests that there is neither bunching at the cut-off with respect to the number of observations per month nor systematic differences in the predetermined covariates around the cut-off. The comparison of the covariates' means around the cut-off in Panel 1 of Table 1 and the results of regressions that use the predetermined covariates as outcome variables in Table 2 confirm the absence of systematic differences around the cut-off. Both tests show that differences in covariates around the cut-off are generally small in size and not significant.[14] There are slightly significant differences in the share of mothers with higher secondary school degrees and fathers with lower secondary school degrees in the two-month window specification with covariates. However, these differences are significant only at the 10% significance level and non-significant in all other specifications. Dustmann et al. (2016) analyse parental characteristics around the school entry cut-off in the German Microcensus 2005 and do not find significant differences. Overall, the evidence suggests that the identifying assumptions hold.

4 Descriptive Statistics

Table 1 presents the descriptive statistics of the sample for respondents who were born before and after the school entry cut-off. The variables' means and standard deviations

[12] Staiger and Stock (1997) suggest that an F-statistic of larger than 10 suffices.

[13] For example, Angrist and Krueger (1991) and Robertson (2011) used season of birth as an instrument.

[14] Moreover, Table A-4 shows that a father's and mother's age and occupational prestige are also balanced around the cut-off.

are shown for the two-month and four-month window. The descriptive statistics of the covariates are reported in Panel 1 and those of the outcomes are in Panel 2.

Focusing on respondents born within the two-month window, the mean values of school starting age show that respondents to the right of the cut-off are, on average, three months older than respondents to the left of the cut-off. The age difference is also illustrated in the upper left graph in Figure 3. Thus, the actual mean difference in school starting age is three months and not eleven months, which would be expected if all children complied with the school entry rule.

About 52% of the respondents in the sample are female and roughly 13% have some migration background. The respondents' fathers are more likely to have a higher secondary school degree than their mothers. Compared to 17% of fathers, only 11% mothers have a high secondary school degree. However, 26% of the mothers hold an intermediate secondary school degree compared to the 21% of fathers. The share of mothers and fathers with either a low or no secondary school degree is roughly the same (59% mothers and 58% fathers).[15]

Panel 2 in Table 1 shows the mean differences in the outcome variables. The sample size in the main analysis varies from 1,674 to 1,890 in the two-month window and 3,391 to 3,856 in the four-month window. In the robustness checks, the sample size varies between 10,400 and 11,784 when all months are included in the estimation.

The descriptive statistics of the two-month window show that the share of smokers among respondents born before the cut-off is 32.4% and thus, 4.3 percentage points higher than that of smokers among respondents born after the cut-off. This difference is almost significant at the 5% significance level with a p-value of 0.051.

Furthermore, respondents born before the cut-off report, on average, significantly lower health than those born after the cut-off. The absolute difference of about 0.138 is significant at the 1% level and about 15% of the variable's standard deviation. Furthermore, the share of persons who report being in good or very good health is about 5.4 percentage points lower among respondents born before the cut-off than those born after; this difference is significant at the 5% level. The quasi-objective SF12 health measures show that the mental health score does not significantly differ between both groups; however, respondents born before the cut-off have a significantly lower physical health score on average than those born after. In terms of one standard deviation, the difference in the physical health score between both groups is about 11.6%.[16] The results for the four-month window confirm the results for the two-month window.

The descriptive statistics of the outcomes imply that individuals born before the school

[15]The values for school degree type do not aggregate to 100% because some respondents' parents have other or unspecified types of school degrees.

[16]The difference in the physical health score is divided by 10, which is the variable's standard deviation in the initial calibration of the SF12 score in the 2004 SOEP sample.

entry cut-off are more likely to smoke and have worse health outcomes than those born after the school entry cut-off. We consider these descriptive results to be rather informative because they resemble an unconditional reduced form estimate for the impact of the distance between the birth month and school entry cut-off on smoking behaviour and health.

5 Results

5.1 Main Results

Regressing smoking behaviour and health outcomes on school starting age using OLS results in non-significant point estimates that are close to zero (Table A-3). Thus, the OLS results suggest that school entry age has no long-term impact on smoking behaviour and health. However, OLS estimation does not take into account that parents strategically enrol their children in school with respect to factors unobserved in the data, and therefore, yields biased estimates. For instance, parents of relatively precocious and independent children might enrol them early and parents of relatively underdeveloped children might enrol them late. Consequently, both moving up and postponing school entry are likely to bias the OLS estimates towards zero.

The following results based on the fuzzy regression discontinuity design account for the endogeneity of school starting age and show that school starting age has sizeable and significant effects on smoking behaviour and health. Table 3 presents the main results of the fuzzy regression discontinuity design. It shows the estimates of the causal effect of a one-month increase in school starting age on the outcomes for the two- and four-month window for three specifications. The first specification includes school starting age as a sole covariate in the regression. The second specification includes indicators for the respondent's gender, birth year, federal state of school entry, and survey year. The third specification comprises indicators for migration background and parental education. The results are robust across specifications: the coefficients' magnitudes and significance report negligible changes. Our preferred specification is the two-month window including all covariates because it most convincingly ensures that persons to the left and right of the cut-off are comparable.

The preferred specification shows that a one-month increase in school starting age decreases the risk of smoking later by about 1.3 percentage points (4.0%). This effect is significant at the 5% level. Moreover, a one-month increase in school starting age increases respondents' health status. The coefficient for the effect of school starting age on the self-reported health scale is 0.042 and significant at the 1% level. The effect corresponds to about 4.5% of one standard deviation. Complementary, the likelihood to report at least good health increases by about 1.6 percentage points (2.4%) and is significant at the 5% level.

The health effect is driven by physical and not mental health. The coefficients for the effect of school starting age on mental health are non-significant.[17] The physical health score, however, significantly increases with school starting age; the coefficient of 0.364 corresponds to about 3.6% of one standard deviation. The results of the four-month window confirm the results of the two-month window. The coefficients have the same sign, are smaller in magnitude but still sizeable and are of similar statistical significance.

5.2 Robustness

The computation of school starting age sa_i and the instrument $older_i$ for respondents without direct information for year, federal state, and school starting age might create measurement error in sa_i and $older_i$. We account for the measurement error in school starting age sa_i using our implemented instrumental variable approach, where sa_i is instrumented by $older_i$.

The potential measurement error in the instrument $older_i$ might be more problematic. Determining a respondent's relevant school entry cut-off by using information on both the federal state and year of the latest school degree might create measurement error in $older_i$ if a respondent relocated across states with different cut-offs between the start of primary school and the completion of secondary school. In the Federal Republic of Germany, school entry cut-offs differed before their harmonisation in 1964. Nonetheless, many federal states shared the same cut-off before harmonisation anyway.

The sub-sample of young respondents with direct information provides information on the extent of mobility during school: only 3.6% of the respondents reported to have moved across federal states between the start of primary school and completion of secondary school. Moreover, the share of respondents who started school in the Federal Republic of Germany before the harmonisation is 17.1%. By multiplying both shares, we estimate that only 0.6% of respondents are misclassified in our sample. However, this figure is likely to be an upper bound because mobility should be lower among older cohorts than among younger ones. Moreover, not every mover relocated from one federal state to another with a different cut-off. Because the risk of misclassifying persons is low, we include respondents at risk of being misclassified to obtain statistical power. Nonetheless, the exclusion of these respondents in Table 4 gives us similar results.

In the main analysis, we use the first available observation for each respondent. Column 2 in Table 5 includes all available observations for each respondent in the es-

[17]Black et al. (2011) find that school starting age has a significant, but small effect on the mental health of 18–20-year-old males. By contrast, we show that school starting age has no significant effect later in life by including both males and females in the analysis.

timation for the main specification, which restricts the sample to a two-month window around the cut-off.[18] The standard errors are clustered at the respondent level. In comparison with the main results in column 1, the significance and magnitudes of the coefficients remain largely unchanged.

Furthermore, columns 3–8 in Table 5 analyse the sensitivity of the main results by including only certain age ranges in the estimation. To check whether the main results are driven by young respondents in the sample, columns 3 and 4 include observations for those aged 30 years and above and columns 5 and 6 include only observations for those 40 years or older. Columns 7 and 8 include observations for those younger than 60 years to avoid the potential effects of retirement on smoking behaviour and health. The point estimate for the effect of school starting age on smoking behaviour is hardly affected and remains significant in all specifications. Although some point estimates for the effect on self-rated health decrease in size and become non-significant for specifications using only the first observation per respondent, the corresponding point estimates using all observations per respondent remain significant.[19] Similarly, the point estimates and significances of school starting age on physical and mental health are barely affected.[20] Overall, the results of Table 5 confirm the main results.

Moreover, Table 6 shows several alternative specifications that include all available months in the estimation, instead of restricting the analysis to two- or four-month windows. In addition to school starting age, the specifications include 1) no further covariates 2) one linear trend in the running variable (with and without covariates) and 3) separate linear trends in the running variable on both sides of the cut-off (with and without covariates).

The first stage equation with separate linear trends in the running variable and covariates is as follows:

$$sa_i = \alpha_0 + \alpha_1 older_i + \alpha_2 dist_i + \alpha_3 dist_i \cdot older_i + \pi X_i' + \gamma_b + \gamma_s + \gamma_w + e_i.$$

The corresponding second stage equation is

$$y_i = \beta_0 + \beta_1 \hat{sa}_i + \beta_2 dist_i + \beta_3 dist_i \cdot older_i + \delta X_i' + \gamma_b + \gamma_s + \gamma_w + u_i.$$

The running variable $dist_i$ denotes the distance between a respondent's month of birth and the school entry cut-off; it is measured in months and takes on integer values between

[18]The results of restricting the sample to a four-month window around the cut-off are shown in Table A-5.

[19]Note that the loss of significance is not surprising given the substantial decrease in the sample size.

[20]Two point estimates for the effect on mental health are significant at the 10% level when young respondents are excluded from the estimation; however, this effect is non-significant when persons older than 60 years are excluded. Moreover, the effect is always non-significant when a four-month window is used instead of a two-month window (see Table A-5).

-5 and 6. Note that the inclusion of quadratic trends would be problematic in the context of this analysis because the running variable has a small number of values. The results in Table 6 are in line with the main results. Overall, the various robustness analyses confirm the main results.

Next, we address the degree of representativeness of the causal long-term effects. The implemented fuzzy regression discontinuity design identifies the local average treatment effect, which is the causal effect for the subgroup of compliers, that is, persons who change their behaviour in compliance with the school entry rule. Table 7 shows that 36% of our sample and 40% of the treated respondents are compliers.[21] The ratio of the likelihood that a complier has a certain characteristic and the general likelihood that a respondent has the same characteristic is close to one for the analysis' predetermined covariates. Thus, the group of compliers is similar to the entire sample with respect to the analysis' predetermined covariates. This similarity indicates that the estimated local average treatment effect could be representative of the entire sample.

6 Mechanisms

In this section, we investigate potential mechanisms through which school starting age might affect smoking behaviour and health. First, drawing on the SOEP, we show suggestive evidence that school starting age affects smoking behaviour and health through peers' age composition. Second, we discuss several studies on Germany that show that school starting age affects children's likelihood to enter a higher secondary school track and thus, their school environment. In addition, we analyse the importance of the school environment mechanism for the effect of school starting age on smoking and health by including the respondent's school degree as a covariate in the regression. Third, we review studies analysing the effect of school starting age on both grade retention and academic achievement because retained students might experience more stress and mental strain, and thus, are more likely to smoke. Fourth, we discuss the results of studies that analyse the effects of school starting age on labour market outcomes.

Peer effects are likely to be an important mechanism because school starting age affects the relative age of school peers. Manski (1993, 1995) points out that it is difficult to disentangle peer effects on individual behaviour into 1) direct effects of peer behaviour (endogenous effect) 2) effects of observed peer characteristics (contextual effect) and 3) effects of unobserved peer characteristics (correlated effect).[22] Most peer effect studies are

[21]To characterise compliers relative to the entire sample, we adopted the methodology as explained in Angrist and Pischke (2009).

[22]For instance, the smoking behaviour of a person's reference group might affect his/her own smoking behaviour (endogenous effect). Moreover, an individual's smoking behaviour may be influenced by the observed socioeconomic status of the reference group (contextual effect). However, it might also be

unable to distinguish between these effects, despite the availability of exogenous variation in peer measures (Sacerdote, 2011). While the effect of school starting age on the relative age of school peers is mainly an endogenous effect, other mechanisms discussed in this section are a combination of endogenous, contextual, and correlated effects.

First, school starting age affects the relative age composition of peers and therefore, the exposure to peer smoking in school. This mechanism can arguably be considered an endogenous effect of peer behaviour. Figure 1 shows that the share of smokers tremendously increases with age during childhood and adolescence. While the share of smokers is close to 0% among 10–11-year-olds, it steadily increases with age to over 40% among 16–17-year-olds. Consequently, children who start school relatively young have peers and friends in school who are both older and more likely to smoke. Thus, analysing the effects of school starting age on the relative age composition of respondents' friends is indicative of the degree of peer smoking in school.

Table 8 shows fuzzy regression discontinuity results in which the characteristics of the respondent's network of friends measured in adulthood—i.e. the network of friends at the time of the SOEP interview—are regressed on school starting age. The estimates are an indication of the impact of school starting age on the characteristics of friends in school under the assumption that childhood friendships persist into adulthood. Whereas the number of friends is unaffected, both the average age and relative age of friends are significantly affected by school starting age.[23] Individuals who started school relatively young are more likely to have older friends later in life and therefore, have increased exposure to smoking in school through older classmates and friends.[24]

Second, school starting age affects a child's likelihood to attend specific school types in secondary education in Germany. Jürges and Schneider (2007), Mühlenweg and Puhani (2010) and Dustmann et al. (2016) find that students who are relatively young at the start of primary school are less likely to attend higher secondary school tracks. Figure 2 shows that the share of smokers in low-track schools is about 5 percentage points higher than medium-track schools and about 11 percentage points higher than high-track ones. Students in low-track schools are therefore, more exposed to peer smoking.

Moreover, students in low-track schools are subject to worse contextual and correlated school and background characteristics than those in higher track schools. According to the German Health Interview and Examination Survey for Children and Adolescents (Hölling et al., 2012), 8.6% of students in low-track schools have at least one parent with an upper secondary school degree compared to 45.7% students in high-track schools. Furthermore,

affected by the unobserved work environment that both the person and reference group share (correlated effect).

[23]The relative age of friends is calculated by dividing the average age of friends by the respondent's own age.

[24]The results are robust to the use of all available months in the estimation with linear trends (see Table A-6).

Dustmann et al. (2016) show that the number of hours taught, teaching intensity and learning goals considerably differ between school tracks. Jürges et al. (2011) show that the increase in the number of high-track schools in post-war Western Germany reduced the rate of smokers through an increase in education—a result that highlights the importance of both years of education and school environment on smoking behaviour.

In Table 9, we include the respondent's highest secondary school degree as a covariate in the estimation to gauge the importance of school environments as a mechanism.[25] In the two-month window, the effect of school starting age on smoking becomes non-significant. Although the point estimate remains sizeable, it decreases in absolute size from -0.013 to -0.009. Moreover, the effect on smoking remains significant in the four-month window. The effect of school starting age on physical health becomes non-significant in the four-month window, but remains significant in the two-month window. The effects on both self-rated health measures remain significant but those on mental health remain non-significant. The significant point estimates for the main specifications decrease in absolute size between 13% for physical health and 31% for smoking behaviour. Thus, school environments are a relevant mechanism, although they do not appear to be the main mechanism through which school starting age affects smoking behaviour and health.

Third, school starting age might affect the likelihood of grade retention. Eide and Showalter (2001), Elder and Lubotsky (2009) and Bernardi (2014) find that an increase in kindergarten or school starting age lowers the risk of grade retention in the US and France. However, Elder and Lubotsky (2009) show that kindergarten starting age increases the likelihood of grade retention mainly in the first and second grade.[26] For Germany, Fertig and Kluve (2005) find that school starting age has no effect on the likelihood of an individual repeating a grade in school.

Thus, school starting age affects grade retention only through the likelihood of repeating a grade in primary school. In contrast, students begin smoking in secondary school, as shown in Figure 1. This result implies that grade retention is an unlikely mechanism through which school starting age affects the likelihood to smoke. In fact, grade retention should lower the risk of smoking among young school starters because it increases their relative age.

Fourth, school starting age might affect smoking behaviour and health through labour market outcomes. Black et al. (2011) find that an increase in school starting age lowers short-term earnings in Norway; however, this effect disappears by the age of 30. Fredriksson and Öckert (2014) show that school starting age affects the timing of labour supply,

[25]For respondents who had not yet finished their secondary education, we included their current school type as a covariate.

[26]Elder and Lubotsky (2009) reveal that a one-year increase in kindergarten entry age decreases the likelihood of grade retention by 13.1 percentage points in the first and second grade and by 15.5 percentage points in any grade in the first eight years of schooling.

but not prime-age earnings in Sweden. Dobkin and Ferreira (2010) find for California and Texas that school starting age has no impact on wages and employment probability. Dustmann et al. (2016) show that school starting age affects the likelihood of students of attending a specific type of secondary school track in Germany; however, there are no long-term effects of tracking on wages, labour force participation, unemployment and occupational choice. They attribute the absence of labour market effects to the flexibility of the German education system, which mitigates mistracking of students.

Overall, school starting age is likely to affect smoking behaviour and health through the relative age composition of peers in school and school environment. By contrast, grade retention is unlikely to increase the risk of smoking because it is affected by school starting age in early grades, while the incidence of smoking occurs in later grades. Labour market outcomes is also an unlikely mechanism through which school starting age affects long-term smoking behaviour and health: while there is some evidence of marginal short-term effects, these effects rapidly disappear. Thus, peer effects are an important mechanism through which school starting age affects long-term smoking behaviour and health.

7 Conclusion

In this study, we examine the long-term effects of school starting age on smoking behaviour and health. Because parents may decide their children's school starting age strategically while considering, for example, the perceived child's school readiness, we implement a fuzzy regression discontinuity design to account for the endogeneity of school starting age. We use exogenous school entry rules, which are based on children's date of birth, as an instrument for the observed school starting age. Our results show that school starting age affects smoking behaviour and health in the long term.

The fuzzy regression discontinuity design results show that a one-month increase in school starting age significantly reduces the long-term risk of smoking by about 1.3 percentage points and increases the long-term likelihood of reporting good or very good health by about 1.6 percentage points. These estimates imply that an increase in school starting age by 11 months—i.e. comparing children who are born in consecutive months around the cut-off and comply with the school entry rule—reduces the risk of smoking by 14.3 percentage points and increases the likelihood of reporting at least good health by about 17.6 percentage points. In addition, controlling for the endogeneity of school starting age is important: the effects of school starting age on smoking behaviour and health estimated using OLS are severely biased towards zero and are non-significant.

Our study shows that the short-term effect of school starting age on adolescent smoking found in previous studies persists into adulthood. Furthermore, our results are qualitatively in line with the literature on the effects of peers on smoking. The results are, however, difficult to quantitatively compare owing to methodological differences. Argys and Rees (2008) find in their preferred OLS specification that female adolescents who were relatively young at school start are 4.1 percentage points more likely to smoke than female adolescents who were relatively old at school start. The point estimates from instrumental variable regressions, although non-significant, imply that both males and females who were young at school start are 1 percentage point more likely to smoke. Studies that adopt school-based peer smoking measures to study the effects of peer behaviour on adolescent smoking find large short-term effects. For instance, Gaviria and Raphael (2001) find that moving a high-school student from a school where no children smoke to one where 25% of children smoke increases the student's likelihood of smoking by 4 percentage points. Powell et al. (2005) show a much larger effect of 14.5 percentage points for the same thought experiment.

We further find that adults who started school relatively young are not only more likely to smoke but also less healthy. This health effect is driven by physical rather than mental health. Elder and Lubotsky (2009), Elder (2010), Evans et al. (2010) and Schwandt and Wuppermann (2016) find that children and adolescents who started school younger have a higher likelihood to be diagnosed with ADHD.[27] Black et al. (2011) report that 18-year-old Norwegian conscripts who started school at a young age are diagnosed with slightly worse mental health in their military medical inspection than those who began school at an older age. While these studies find short-term effects of school starting age on mental health measures, we do not find significant long-term effects.

Similar to previous studies on peer effects, the causal estimates in this study are a combined effect of peer behaviour, peer characteristics and school environment. We analyse and discuss important channels through which school starting age might affect smoking behaviour and health to shed light on the mechanisms underlying our estimates.

First, students who start school young have older class and school peers. These young students are exposed to peer smoking earlier than those who start school when they are older, because the prevalence of smoking considerably increases with age during childhood and adolescence. We show that adults who started school relatively young have, on average, older friends than those who started school older. We interpret this result as suggestive evidence that younger students are influenced by older peers and friends.

Second, Jürges and Schneider (2007), Mühlenweg and Puhani (2010) and Dustmann et al. (2016) show that students in Germany who start primary school relatively young are less likely to attend a higher secondary school track. At the same time, higher track schools have lower shares of smokers and more school resources. We show that school environment partially explains the effects of school starting age on smoking behaviour and health. School environment does, however, explain only a small share of the effects

[27]These studies interpret this finding as evidence for misdiagnosis of ADHD.

of school starting age. Furthermore, we discuss that grade retention and labour market outcomes are unlikely mechanisms through which school starting age affects smoking behaviour and health in the long term.

References

Andersen, H. H., A. Mühlbacher, M. Nübling, J. Schupp, and G. G. Wagner (2007): "Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF-12v2," *Schmollers Jahrbuch*, 127, 171–182.

Anderson, P. M., K. F. Butcher, E. U. Cascio, and D. W. Schanzenbach (2011): "Is Being in School Better? The Impact of School on Children's BMI When Starting Age Is Endogenous," *Journal of Health Economics*, 30, 977–986.

Anger, S., M. Kvasnicka, and T. Siedler (2011): "One Last Puff? Public Smoking Bans and Smoking Behavior," *Journal of Health Economics*, 30, 591–601.

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

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

Argys, L. M. and D. I. Rees (2008): "Searching for Peer Group Effects: A Test of the Contagion Hypothesis," *Review of Economics and Statistics*, 90, 442–458.

Bedard, K. and E. Dhuey (2006): "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *Quarterly Journal of Economics*, 121, 1437–1472.

Bernardi, F. (2014): "Compensatory Advantage as a Mechanism of Educational Inequality: A Regression Discontinuity Based on Month of Birth," *Sociology of Education*, 87, 74–88.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011): "Too Young to Leave the Nest? The Effects of School Starting Age," *Review of Economics and Statistics*, 93, 455–467.

Chassin, L., C. C. Presson, J. S. Rose, and S. J. Sherman (1996): "The Natural History of Cigarette Smoking from Adolescence to Adulthood: Demographic Predictors of Continuity and Change," *Health Psychology*, 15, 478–484.

Dalsgaard, S., M. K. Humlum, H. S. Nielsen, and M. Simonsen (2012): "Relative Standards in ADHD Diagnoses: The Role of Specialist Behavior," *Economics Letters*, 117, 663–665.

Destatis (2014): "Mikrozensus - Fragen zur Gesundheit - Rauchgewohnheiten der Bevölkerung," Publikationen im Bereich Gesundheitszustand, Statistisches Bundesamt, Wiesbaden.

Dobkin, C. and F. Ferreira (2010): "Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes?" *Economics of Education Review*, 29, 40–54.

Dustmann, C., P. A. Puhani, and U. Schönberg (2016): "The Long-Term Effects of Early Track Choice," *Economic Journal*, forthcoming.

Eibich, P. (2015): "Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity," *Journal of Health Economics*, 43, 1–12.

Eide, E. R. and M. H. Showalter (2001): "The Effect of Grade Retention on Educational and Labor Market Outcomes," *Economics of Education Review*, 20, 563–576.

Elder, T. E. (2010): "The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates," *Journal of Health Economics*, 29, 641–656.

Elder, T. E. and D. H. Lubotsky (2009): "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, 44, 641–683.

Evans, W. N., M. S. Morrill, and S. T. Parente (2010): "Measuring Inappropriate Medical Diagnosis and Treatment in Survey Data: The Case of ADHD among School-Age Children," *Journal of Health Economics*, 29, 657–673.

Fertig, M. and J. Kluve (2005): "The Effect of Age at School Entry on Educational Achievement in Germany," IZA Discussion Paper Series 1507.

Fletcher, J. and T. Kim (2016): "The Effects of Changes in Kindergarten Entry Age Policies on Educational Achievement," *Economics of Education Review*, 50, 45–62.

Fredriksson, P. and B. Öckert (2014): "Life-Cycle Effects of Age at School Start," *Economic Journal*, 124, 977–1004.

Gaviria, A. and S. Raphael (2001): "School-Based Peer Effects and Juvenile Behavior," *Review of Economics and Statistics*, 83, 257–268.

Godard, M. (2016): "Gaining Weight through Retirement? Results from the SHARE Survey," *Journal of Health Economics*, 45, 27–46.

Graue, M. E. and J. DiPerna (2000): "Redshirting and Early Retention: Who Gets the 'Gift of Time' and What Are Its Outcomes?" *American Educational Research Journal*, 37, 509–534.

Gruber, J. and J. Zinman (2000): "Youth Smoking in the U.S.: Evidence and Implications," Working Paper 7780, National Bureau of Economic Research.

Hahn, J., P. Todd, and W. V. d. Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69, 201–209.

Hölling, H., R. Schlack, P. Kamtsiuris, H. Butschalowsky, M. Schlaud, and B. Kurth (2012): "Die KiGGS-Studie: Bundesweit repräsentative Längs- und Querschnittstudie zur Gesundheit von Kindern und Jugendlichen im Rahmen des Gesundheitsmonitorings am Robert Koch-Institut," *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 55, 836–842.

Imbens, G. W. and T. Lemieux (2008): "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142, 615–635.

Jürges, H., S. Reinhold, and M. Salm (2011): "Does Schooling Affect Health Behavior? Evidence from the Educational Expansion in Western Germany," *Economics of Education Review*, 30, 862–872.

Jürges, H. and K. Schneider (2007): "What Can Go Wrong Will Go Wrong: Birthday Effects and Early Tracking in the German School System," CESifo Working Paper 2055.

Landersø, R., H. S. Nielsen, and M. Simonsen (2016): "School Starting Age and the Crime-Age Profile," *Economic Journal*, forthcoming.

Lee, D. S. and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48, 281–355.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

——— (1995): *Identification Problems in the Social Sciences*, Cambridge, Mass.: Harvard University Press.

Marcus, J. (2013): "The Effect of Unemployment on the Mental Health of Spouses – Evidence from Plant Closures in Germany," *Journal of Health Economics*, 32, 546–558.

McEwan, P. J. and J. S. Shapiro (2008): "The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates," *The Journal of Human Resources*, 43, 1–29.

Morrow, R. L., E. J. Garland, J. M. Wright, M. Maclure, S. Taylor, and C. R. Dormuth (2012): "Influence of Relative Age on Diagnosis and Treatment of Attention-Deficit/Hyperactivity Disorder in Children," *Canadian Medical Association Journal*, 184, 755–762.

Mühlenweg, A., D. Blomeyer, H. Stichnoth, and M. Laucht (2012): "Effects of Age at School Entry (ASE) on the Development of Non-cognitive Skills: Evidence from Psychometric Data," *Economics of Education Review*, 31, 68–76.

Mühlenweg, A. M. and P. A. Puhani (2010): "The Evolution of the School-Entry Age Effect in a School Tracking System," *Journal of Human Resources*, 45, 407–438.

Norton, E. C., R. C. Lindrooth, and S. T. Ennett (1998): "Controlling for the Endogeneity of Peer Substance Use on Adolescent Alcohol and Tobacco Use," *Health Economics*, 7, 439–453.

Powell, L. M., J. A. Tauras, and H. Ross (2005): "The Importance of Peer Effects, Cigarette Prices and Tobacco Control Policies for Youth Smoking Behavior," *Journal of Health Economics*, 24, 950–968.

Puhani, P. A. and A. M. Weber (2007): "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Early Educational Effects of Age of School Entry in Germany," *Empirical Economics*, 32, 359–386.

Robertson, E. (2011): "The Effects of Quarter of Birth on Academic Outcomes at the Elementary School Level," *Economics of Education Review*, 30, 300–311.

Sacerdote, B. (2011): "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" in *Handbook of the Economics of Education*, ed. by E. A. Hanushek, S. Machin, and L. Woessmann, Elsevier, vol. 3, 249–277.

Salyers, M. P., H. B. Bosworth, J. W. Swanson, J. Lamb-Pagone, and F. C. Osher (2000): "Reliability and Validity of the SF-12 Health Survey Among People With Severe Mental Illness." *Medical Care*, 38, 1141–1150.

Schwandt, H. and A. Wuppermann (2016): "The Youngest Get the Pill: ADHD Misdiagnosis in Germany, Its Regional Correlates and International Comparison," *Labour Economics*, forthcoming.

Staiger, D. and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Stipek, D. (2002): "At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents," SRCD Social Policy Report 16(2).

U.S. Department of Health and Human Services (2012): "Preventive Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General," Report, Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA.

Wacker, M., R. Holle, J. Heinrich, K.-H. Ladwig, A. Peters, R. Leidl, and P. Menn (2013): "The Association of Smoking Status with Healthcare Utilisation, Productivity Loss and Resulting Costs: Results from the Population-Based KORA F4 Study," *BMC Health Services Research*, 13, 278.

Wagner, G. G., J. R. Frick, and J. Schupp (2007): "The German Socio-economic Panel Study (SEOP) — Scope, Evolution and Enhancements," *Schmollers Jahrbuch: Zeitschrift fur Wirtschafts- und Sozialwissenschaften/Journal of Applied Social Science Studies*, 127, 139–169.

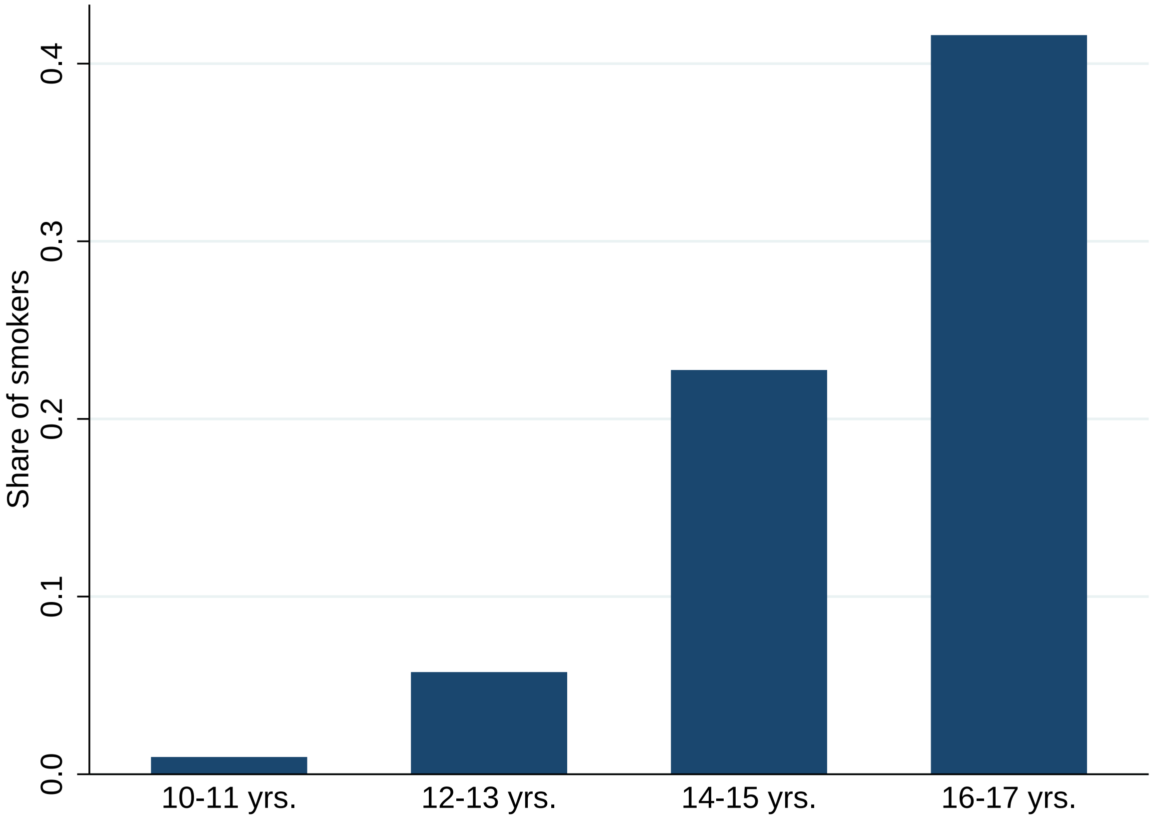
Ware, Jr, J., M. Kosinski, and S. D. Keller (1996): "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity," *Medical Care*, 34, 220–233.

WHO (2015a): "WHO Report on the Global Tobacco Epidemic, 2015: Raising Taxes on Tobacco," World Health Organization.

——— (2015b): "WHO Report on the Global Tobacco Epidemic, 2015: Raising Taxes on Tobacco, Appendix X," World Health Organization.

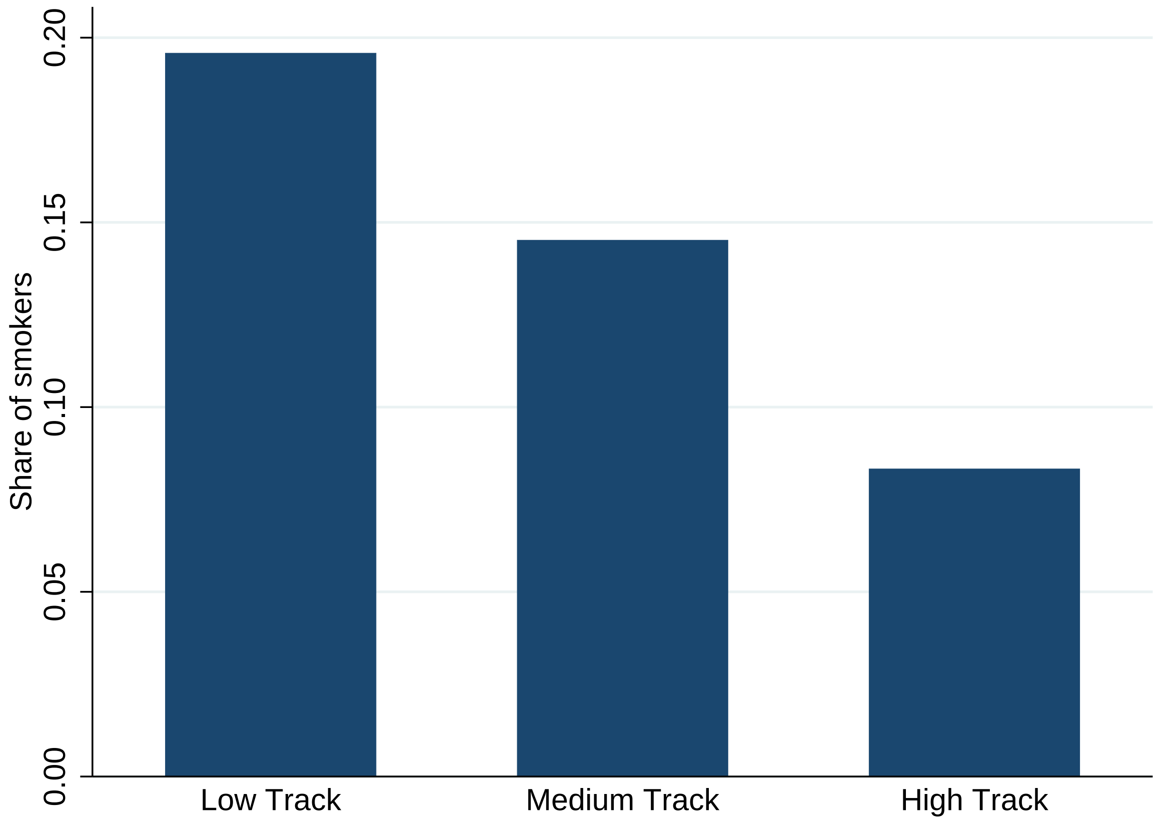
Xu, X., E. E. Bishop, S. M. Kennedy, S. A. Simpson, and T. F. Pechacek (2015): "Annual Healthcare Spending Attributable to Cigarette Smoking: An Update," *American Journal of Preventive Medicine*, 48, 326–333.

Figure 1: Share of smokers by age.



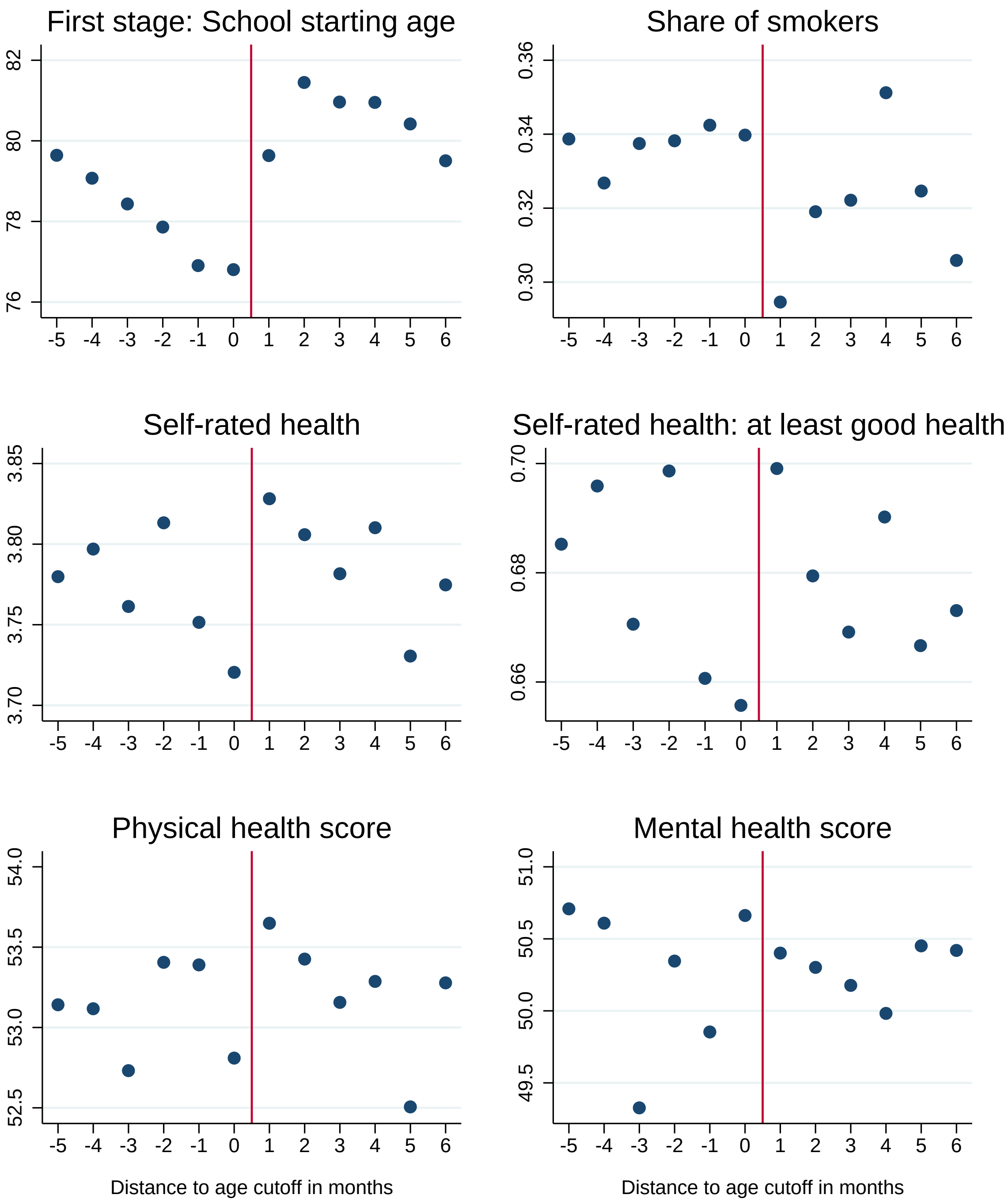
Notes: The figure shows a bar plot of the share of smokers by age. *Source*: German Health Interview and Examination Survey for Children and Adolescents (KiGGS) 2003–2006.

Figure 2: Share of smokers by type of secondary school.



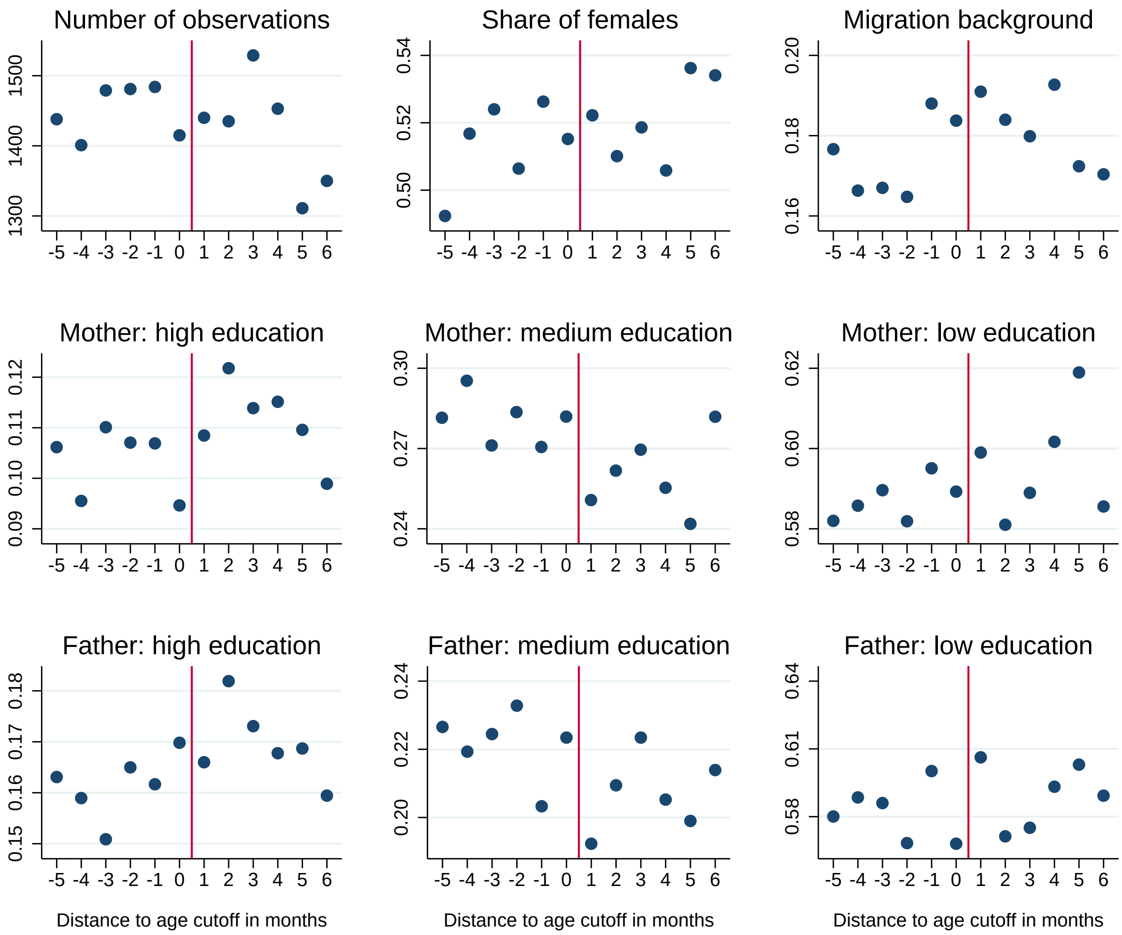
Notes: The figure shows a bar plot of the share of smokers among 12–15-year-olds for the three main secondary school types in Germany. 'Low track' refers to low-track schools (*Hauptschule*), 'medium track' is medium-track schools (*Realschule*) and 'high track' is upper-track schools (*Gymnasium*). *Source*: German Health Interview and Examination Survey for Children and Adolescents (KiGGS), 2003–2006.

Figure 3: Means of school starting age and outcomes.



Notes: The graphs show the average value of the variables for each value of the running variable. The running variable is defined as the distance between a respondent's birth month and the relevant school entry cut-off. The cut-off month, which includes the cut-off day, is assigned the value of zero. The cut-off day is always the last day of the respective cut-off month (e.g. June 30th). *Source*: German Socio-Economic Panel (SOEP).

Figure 4: Number of observations and means of covariates.



Notes: The graphs show the average value of the number of observations or the respective covariate for each value of the running variable. The running variable is defined as the distance between a respondent's birth month and the relevant school entry cut-off. The cut-off month, which includes the cut-off day, is assigned the value of zero. The cut-off day is always the last day of the respective cut-off month (e.g. June 30th). *Source*: German Socio-Economic Panel (SOEP).

Table 1: Descriptive statistics.

	2 months			4 months		
	Before	After	p-value	Before	After	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel 1: Covariates					
School starting age	76.894	80.068	0.000	76.939	80.936	0.000
	(7.650)	(8.103)		(7.356)	(8.030)	
Female	0.523	0.532	0.683	0.524	0.529	0.755
	(0.500)	(0.499)		(0.500)	(0.499)	
Migration background	0.118	0.143	0.102	0.123	0.137	0.186
	(0.323)	(0.351)		(0.329)	(0.344)	
Highest school degree: mother						
High	0.097	0.110	0.340	0.101	0.116	0.134
	(0.296)	(0.313)		(0.302)	(0.321)	
Medium	0.281	0.249	0.116	0.275	0.254	0.138
	0.450	0.433		0.447	0.435	
Low	0.589	0.597	0.708	0.593	0.592	0.917
	(0.492)	(0.491)		(0.491)	(0.492)	
Highest school degree: father						
High	0.171	0.169	0.946	0.168	0.177	0.443
	(0.376)	(0.375)		(0.374)	(0.382)	
Medium	0.224	0.195	0.111	0.214	0.202	0.331
	(0.417)	(0.396)		(0.411)	(0.401)	
Low	0.565	0.599	0.129	0.582	0.585	0.884
	(0.496)	(0.490)		(0.493)	(0.493)	
n	973	956		1,982	1,952	

continued on next page ...

... continued from previous page

	2 months			4 months		
	Before	After	p-value	Before	After	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel 2: Outcomes			
Smoking	0.324	0.281	0.051	0.331	0.294	0.019
	(0.468)	(0.450)		(0.471)	(0.456)	
n	866	833		1,765	1,684	
Self-rated health	3.720	3.858	0.001	3.739	3.844	0.000
	(0.935)	(0.889)		(0.944)	(0.885)	
Self-rated health:	0.654	0.708	0.013	0.659	0.698	0.010
at least good	(0.476)	(0.455)		(0.474)	(0.459)	
n	952	938		1,948	1,908	
SF12: Physical health	52.840	53.997	0.006	53.151	53.755	0.043
	(9.244)	(8.034)		(8.961)	(8.388)	
SF12: Mental health	50.776	50.466	0.518	50.482	50.471	0.974
	(9.872)	(9.743)		(9.814)	(9.638)	
n	853	821		1,740	1,651	
Number of friends	4.865	4.788	0.714	5.172	4.700	0.168
	(3.857)	(4.326)		(13.017)	(3.956)	
n	770	770		1,580	1,566	
Average age of friends	37.196	35.715	0.111	36.899	35.701	0.071
	(12.906)	(13.006)		(12.950)	(13.198)	
Relative age of friends	1.038	1.012	0.055	1.036	1.019	0.125
	(0.220)	(0.164)		(0.218)	(0.214)	
n	390	390		800	753	

Notes: The table shows the descriptive statistics of the estimation sample around the cut-off for the two- and four-months window. The variables' means and standard deviations are in parentheses. Columns 3 and 6 report the p-values of the simple t-tests for the differences in the variables' means before and after the school entry cut-off. Note that the sample size 'n' varies across outcomes in Panel 2, because information on certain outcomes is not available for each respondent and survey wave.

Table 2: Instrument validity: pretreatment covariates as dependent variables.

	2 months		4 months	
	(1)	(2)	(3)	(4)
	Female			
School starting age	0.003	0.004	0.001	0.001
	(0.007)	(0.007)	(0.004)	(0.004)
	Migration background			
School starting age	0.008	0.005	0.004	0.001
	(0.005)	(0.004)	(0.003)	(0.002)
	Secondary school degree: mother			
	High			
School starting age	0.004	0.006*	0.004	0.002
	(0.004)	(0.004)	(0.003)	(0.002)
	Medium			
School starting age	−0.010	−0.003	−0.005	−0.003
	(0.007)	(0.005)	(0.004)	(0.003)
	Low			
School starting age	0.003	−0.006	−0.000	−0.000
	(0.007)	(0.005)	(0.004)	(0.003)
	Secondary school degree: father			
	High			
School starting age	−0.000	−0.002	0.002	0.000
	(0.005)	(0.004)	(0.003)	(0.002)
	Medium			
School starting age	−0.009	−0.004	−0.003	−0.001
	(0.006)	(0.005)	(0.003)	(0.003)
	Low			
School starting age	0.011	0.010*	0.001	0.001
	(0.007)	(0.005)	(0.004)	(0.003)
n	1,910	1,910	3,864	3,864
Birth year indicators	No	Yes	No	Yes
Federal state indicators	No	Yes	No	Yes
Additional covariates	No	Yes	No	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design, in which predetermined covariates are regressed on school starting age. Additional covariates include female, migration background, highest degree held by both the father and mother. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Fuzzy RDD: smoking behaviour and health.

	2 months			4 months		
	(1)	(2)	(3)	(4)	(5)	(6)
			Smoking behavior			
School starting age	-0.014^{*}	-0.013^{*}	-0.013^{**}	-0.009^{**}	-0.009^{**}	-0.009^{**}
	(0.007)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)
1st stage: F-Statistic	65.7	77.1	78.2	220.6	239.2	246.2
n	1,699	1,699	1,699	3,449	3,449	3,449
			Self-rated health			
School starting age	0.044^{***}	0.043^{***}	0.042^{***}	0.026^{***}	0.025^{***}	0.025^{***}
	(0.014)	(0.013)	(0.013)	(0.008)	(0.007)	(0.007)
1st stage: F-Statistic	74.0	84.2	85.1	254.1	278.3	284.1
n	1,890	1,890	1,890	3,856	3,856	3,856
			Self-rated health: at least good			
School starting age	0.017^{**}	0.017^{**}	0.016^{**}	0.010^{**}	0.009^{**}	0.009^{***}
	(0.007)	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)
1st stage: F-Statistic	74.0	84.2	85.1	254.1	278.3	284.1
n	1,890	1,890	1,890	3,856	3,856	3,856
			SF12: physical health score			
School starting age	0.374^{***}	0.362^{***}	0.364^{***}	0.155^{**}	0.148^{**}	0.148^{**}
	(0.143)	(0.129)	(0.126)	(0.077)	(0.071)	(0.069)
1st stage: F-Statistic	63.8	77.1	79.1	214.4	234.5	242.0
n	1,674	1,674	1,674	3,391	3,391	3,391
			SF12: mental health score			
School starting age	−0.100	−0.139	−0.144	−0.003	0.004	0.007
	(0.155)	(0.141)	(0.139)	(0.086)	(0.082)	(0.081)
1st stage: F-Statistic	63.8	77.1	79.1	214.4	234.5	242.0
n	1,674	1,674	1,674	3,391	3,391	3,391
Birth year indicators	No	Yes	Yes	No	Yes	Yes
Survey year indicators	No	Yes	Yes	No	Yes	Yes
Federal state indicators	No	Yes	Yes	No	Yes	Yes
Female	No	Yes	Yes	No	Yes	Yes
Migration background	No	No	Yes	No	No	Yes
Education father	No	No	Yes	No	No	Yes
Education mother	No	No	Yes	No	No	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Fuzzy RDD: excluding potentially misclassified respondents.

	2 months			4 months		
	(1)	(2)	(3)	(4)	(5)	(6)
			Smoking behavior			
School starting age	-0.016^{*}	-0.015^{*}	-0.016^{*}	-0.010^{**}	-0.009^{**}	-0.009^{**}
	(0.009)	(0.009)	(0.009)	(0.005)	(0.005)	(0.005)
1st stage: F-Statistic	55.0	53.7	54.6	191.3	192.8	199.3
n	1,399	1,399	1,399	2,854	2,854	2,854
			Self-rated health			
School starting age	0.064^{***}	0.060^{***}	0.058^{***}	0.030^{***}	0.027^{***}	0.027^{***}
	(0.018)	(0.017)	(0.016)	(0.008)	(0.008)	(0.008)
1st stage: F-Statistic	60.5	60.7	61.0	224.2	231.5	236.9
n	1,571	1,571	1,571	3,219	3,219	3,219
			Self-rated health: at least good			
School starting age	0.027^{***}	0.024^{***}	0.024^{***}	0.012^{***}	0.011^{***}	0.011^{***}
	(0.009)	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)
1st stage: F-Statistic	60.5	60.7	61.0	224.2	231.5	236.9
n	1,571	1,571	1,571	3,219	3,219	3,219
			SF12: physical health score			
School starting age	0.361^{**}	0.281^{*}	0.283^{*}	0.106	0.059	0.064
	(0.165)	(0.154)	(0.151)	(0.085)	(0.080)	(0.078)
1st stage: F-Statistic	52.2	53.5	55.1	184.4	187.2	194.0
n	1,373	1,373	1,373	2,795	2,795	2,795
			SF12: mental health score			
School starting age	−0.031	0.001	−0.009	0.075	0.107	0.110
	(0.190)	(0.184)	(0.182)	(0.100)	(0.098)	(0.097)
1st stage: F-Statistic	52.2	53.5	55.1	184.4	187.2	194.0
n	1,373	1,373	1,373	2,795	2,795	2,795
Birth year indicators	No	Yes	Yes	No	Yes	Yes
Survey year indicators	No	Yes	Yes	No	Yes	Yes
Federal state indicators	No	Yes	Yes	No	Yes	Yes
Female	No	Yes	Yes	No	Yes	Yes
Migration background	No	No	Yes	No	No	Yes
Education father	No	No	Yes	No	No	Yes
Education mother	No	No	Yes	No	No	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design, excluding respondents who started school prior to 1964 and are thus, at risk of being misclassified with respect to whether they are to the left or right of the cut-off. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Fuzzy RDD including all observations per respondent and sensitivity by age group (two-month window).

	The sample includes							
	All ages		Age ≥ 30		Age ≥ 40		Age < 60	
	1st obs.	All obs.	1st obs.	All obs.	1st obs.	All obs.	1st obs.	All obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Smoking behavior							
School starting age	−0.013**	−0.012*	−0.014*	−0.017**	−0.014*	−0.018**	−0.013*	−0.012*
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
1st stage: F-Statistic	78.2	58.9	49.3	39.2	45.1	35.7	75.5	59.6
n	1,699	5,460	1,114	3,685	783	2,654	1,633	5,291
	Self-rated health							
School starting age	0.042***	0.031***	0.021	0.026**	0.028*	0.028***	0.036***	0.027**
	(0.013)	(0.010)	(0.014)	(0.011)	(0.015)	(0.011)	(0.013)	(0.011)
1st stage: F-Statistic	85.1	63.9	54.8	55.7	50.9	57.2	81.1	60.8
n	1,890	11,014	1,189	9,235	825	8,203	1,822	5,306
	Self-rated health: at least good							
School starting age	0.016**	0.014***	0.007	0.011**	0.010	0.012**	0.014**	0.015***
	(0.006)	(0.005)	(0.007)	(0.005)	(0.008)	(0.005)	(0.007)	(0.006)
1st stage: F-Statistic	85.1	63.9	54.8	55.7	50.9	57.2	81.1	60.8
n	1,890	11,014	1,189	9,235	825	8,203	1,822	5,306
	SF12: physical health score							
School starting age	0.364***	0.321***	0.368**	0.334**	0.506***	0.371**	0.289**	0.294***
	(0.126)	(0.110)	(0.154)	(0.133)	(0.177)	(0.150)	(0.122)	(0.107)
1st stage: F-Statistic	79.1	60.5	55.0	41.1	46.6	33.8	76.4	61.5
n	1,674	4,692	1,115	3,283	797	2,413	1,608	4,528
	SF12: mental health score							
School starting age	−0.144	−0.173	−0.217	−0.244*	−0.212	−0.258*	−0.166	−0.178
	(0.139)	(0.117)	(0.157)	(0.129)	(0.164)	(0.136)	(0.142)	(0.116)
1st stage: F-Statistic	79.1	60.5	55.0	41.1	46.6	33.8	76.4	61.5
n	1,674	4,692	1,115	3,283	797	2,413	1,608	4,528
Birth year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Federal state indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Migration background	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education father	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education mother	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design, restricting the sample to a two-month window around the school entry cut-off. Analogous to the main analysis, '1st obs.' columns include only the first observation per respondent and 'All obs.' columns include all available observations per respondent. Robust standard errors are in parentheses. The models including all waves use clustered standard errors at the level of respondents. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Fuzzy RDD: all months and trends.

	No trend	General linear trend		Seperate linear trends	
	(1)	(2)	(3)	(4)	(5)
	Smoking behaviour				
School starting age	-0.008^{**}	-0.008^{**}	-0.006^{*}	-0.008^{**}	-0.007^{*}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
1st stage: F-Statistic	312.6	246.3	275.2	218.2	243.6
n	10,590	10,590	10,590	10,590	10,590
	Self-rated health				
School starting age	0.013^{*}	0.025^{***}	0.023^{***}	0.027^{***}	0.024^{***}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
1st stage: F-Statistic	333.4	287.7	316.2	256.7	281.5
n	11,784	11,784	11,784	11,784	11,784
	Self-rated health: at least good				
School starting age	0.003	0.009^{***}	0.008^{**}	0.009^{**}	0.008^{**}
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)
1st stage: F-Statistic	333.4	287.7	316.2	256.7	281.5
n	11,784	11,784	11,784	11,784	11,784
	SF12: physical health score				
School starting age	0.076	0.161^{**}	0.147^{**}	0.187^{**}	0.167^{**}
	(0.067)	(0.073)	(0.066)	(0.077)	(0.069)
1st stage: F-Statistic	300.7	238.6	266.8	211.2	235.6
n	10,400	10,400	10,400	10,400	10,400
	SF12: mental health score				
School starting age	0.014	0.021	0.016	0.016	0.012
	(0.074)	(0.081)	(0.078)	(0.086)	(0.082)
1st stage: F-Statistic	300.6	238.1	266.8	210.3	235.6
n	10,400	10,400	10,400	10,400	10,400
Birth year indicators	No	No	Yes	No	Yes
Survey year indicators	No	No	Yes	No	Yes
Federal state indicators	No	No	Yes	No	Yes
Female	No	No	Yes	No	Yes
Education father	No	No	Yes	No	Yes
Education mother	No	No	Yes	No	Yes
Migration background	No	No	Yes	No	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Characterisation of compliers.

	Share
Compliers in sample	0.36
Compliers among treated individuals	0.40

	Ratio
Female	1.03
Highest degree: mother	
High degree	1.13
Medium degree	1.16
Low Degree	0.94
Highest degree: father	
High degree	0.97
Medium degree	1.18
Low Degree	0.99
Migration Background	0.95

Notes: The treatment variable school starting age sa_i is dichotomised. The ratio denotes the likelihood that a complier has a certain characteristic divided by the likelihood that an individual has the same characteristic: $\frac{P[x_i=1|Complier]}{P[x_i=1]}$.

Table 8: Fuzzy RDD: network of friends.

	2 months			4 months		
	(1)	(2)	(3)	(4)	(5)	(6)
			Number of friends			
School starting age	−0.025	−0.005	−0.009	−0.120	−0.107	−0.110
	(0.069)	(0.063)	(0.061)	(0.088)	(0.074)	(0.074)
1st stage: F-Statistic	55.2	63.5	66.6	198.8	215.7	224
n	1,540	1,540	1,540	3,146	3,146	3,146
			Average age of friends			
School starting age	−0.403	−0.256**	−0.243**	−0.309*	−0.192**	−0.193**
	(0.250)	(0.115)	(0.112)	(0.170)	(0.083)	(0.081)
1st stage: F-Statistic	42.9	44.2	45.3	101.1	103.6	108.0
n	780	780	780	1,553	1,553	1,553
			Relative age of friends			
School starting age	−0.008*	−0.008**	−0.008**	−0.005	−0.005	−0.005*
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
1st stage: F-Statistic	42.9	44.2	45.3	101.1	103.6	108.0
n	780	780	780	1,553	1,553	1,553
Birth year indicators	No	Yes	Yes	No	Yes	Yes
Survey year indicators	No	Yes	Yes	No	Yes	Yes
Federal state indicators	No	Yes	Yes	No	Yes	Yes
Female	No	Yes	Yes	No	Yes	Yes
Migration background	No	No	Yes	No	No	Yes
Education father	No	No	Yes	No	No	Yes
Education mother	No	No	Yes	No	No	Yes

Notes: Family members and relatives are excluded. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Fuzzy RDD: secondary school degree as a mechanism.

	2 months		4 months	
	(1)	(2)	(3)	(4)
	Smoking behavior			
School starting age	-0.013^{**}	-0.009	-0.009^{**}	-0.006^{*}
	(0.007)	(0.007)	(0.004)	(0.004)
1st stage: F-Statistic	78.2	75.5	246.2	239.0
n	1,699	1,699	3,449	3,449
	Self-rated health			
School starting age	0.042^{***}	0.034^{***}	0.025^{***}	0.021^{***}
	(0.013)	(0.013)	(0.007)	(0.007)
1st stage: F-Statistic	85.1	83.9	284.1	273.4
n	1,890	1,890	3,856	3,856
	Self-rated health: at least good			
School starting age	0.016^{**}	0.012^{*}	0.009^{***}	0.007^{**}
	(0.006)	(0.006)	(0.004)	(0.004)
1st stage: F-Statistic	85.1	83.9	284.1	273.4
n	1,890	1,890	3,856	3,856
	SF12: physical health score			
School starting age	0.364^{***}	0.317^{**}	0.148^{**}	0.113
	(0.126)	(0.123)	(0.069)	(0.069)
1st stage: F-Statistic	79.1	79.8	242.0	236.9
n	1,674	1,674	3,391	3,391
	SF12: mental health score			
School starting age	-0.144	-0.140	0.007	-0.006
	(0.139)	(0.139)	(0.081)	(0.082)
1st stage: F-Statistic	79.1	79.8	242.0	236.9
n	1,674	1,674	3,391	3,391
Birth year indicators	Yes	Yes	Yes	Yes
Survey year indicators	Yes	Yes	Yes	Yes
Federal state indicators	Yes	Yes	Yes	Yes
Female	Yes	Yes	Yes	Yes
Migration background	Yes	Yes	Yes	Yes
Education father	Yes	Yes	Yes	Yes
Education mother	Yes	Yes	Yes	Yes
Education respondent	No	Yes	No	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design. Columns 2 and 4 include the respondent's highest secondary school degree. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

Table A-1: Average age of respondents by outcome.

	2 months			4 months		
	Before	After	p-value	Before	After	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
			Smoking			
Age at interview	36.8	37.0	0.764	36.8	36.4	0.473
	(13.7)	(14.0)		(13.7)	(14.0)	
n	866	833		1765	1684	
			Self-rated health			
Age at interview	36.1	36.0	0.958	36.1	35.7	0.440
	(13.9)	(14.1)		(13.8)	(14.1)	
n	952	938		1948	1908	
			SF12: physical and mental health scores			
Age at interview	37.4	37.7	0.716	37.4	37.1	0.603
	(13.5)	(13.8)		(13.4)	(13.7)	
n	853	821		1740	1651	
			Number of friends			
Age at interview	37.7	37.9	0.745	37.7	37.6	0.834
	(13.2)	(13.7)		(13.3)	(13.7)	
n	770	770		1580	1566	
			Average and relative age of friends			
Age at interview	36.3	35.7	0.527	36.1	35.5	0.370
	(12.3)	(13.0)		(12.3)	(13.0)	
n	390	390		800	753	

Notes: The table shows the respondents' mean age before and after the school entry cut-off at the time of the SOEP interview by outcome for both the two- and the four-month window around the school entry cut-off. Columns 3 and 6 report the p-values of simple t-tests for the differences before and after the school entry cut-off.

Table A-2: Availability of outcome variables in the SOEP.

Year	Smoking	Self-rated health	SF12	Number of friends	Average and relative age of friends
1992		X			
1994		X			
1995		X			
1996		X			
1997		X			
1998	X	X			
1999	X	X			
2000		X			
2001	X	X			
2002	X	X	X		
2003		X		X	
2004	X	X	X		
2005		X			
2006	X	X	X		X
2007		X			
2008	X	X	X	X	
2009		X			
2010	X	X	X		
2011		X		X	X
2012	X	X	X		
2013		X		X	

Table A-3: OLS and fuzzy RDD results.

	Smoking behavior		Self-rated health		Self-rated health: at least good		Physical health		Mental health	
	2 months	4 months	2 months	4 months	2 months	4 months	2 months	4 months	2 months	4 months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
OLS										
School starting age	−0.001	−0.001	0.001	0.001	0.001	0.000	−0.006	0.008	−0.002	−0.009
	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)	(0.001)	(0.028)	(0.019)	(0.031)	(0.022)
Fuzzy RDD: 2nd stage										
School starting age	−0.013**	−0.009**	0.042***	0.025***	0.016**	0.009***	0.364***	0.148**	−0.144	0.007
	(0.007)	(0.004)	(0.013)	(0.007)	(0.006)	(0.004)	(0.126)	(0.069)	(0.139)	(0.081)
Reduced form										
Older	−0.044*	−0.035**	0.138***	0.103***	0.055**	0.038***	1.235***	0.598**	−0.489	0.027
	(0.023)	(0.016)	(0.041)	(0.028)	(0.021)	(0.015)	(0.413)	(0.282)	(0.478)	(0.333)
Fuzzy RDD: 1st stage										
Older	3.346***	4.043***	3.330***	4.119***	3.330***	4.119***	3.387***	4.053***	3.387***	4.053***
	(0.378)	(0.258)	(0.361)	(0.244)	(0.361)	(0.244)	(0.381)	(0.261)	(0.381)	(0.261)
n	1,699	3,449	1,890	3,856	1,890	3,856	1,674	3,391	1,674	3,391

Notes: The table shows the OLS results, second stage results of the fuzzy RDD, reduced form results and first stage results of the fuzzy RDD. All specifications include indicators for the respondent's gender, migration background and highest school degree held by both the father and mother. They also include indicators for birth year, federal state of school entry and survey year. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A-4: Further parental characteristics at the school entry cutoff.

	2 months			4 months		
	Before	After	p-value	Before	After	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Year of birth: mother	1943.4	1942.8	0.358	1943.1	1943.3	0.686
	(14.9)	(15.0)		(14.8)	(15.2)	
n	962	944		1,958	1,934	
Year of birth: father	1940.3	1940.2	0.792	1940.0	1940.5	0.310
	(15.6)	(15.4)		(15.5)	(15.5)	
n	959	939		1,946	1,922	
Occupational score: mother	58.3	60.5	0.206	59.2	60.5	0.286
	(27.3)	(29.3)		(27.0)	(28.5)	
n	575	542		1,131	1,117	
Occupational score: father	60.7	61.1	0.812	60.6	60.5	0.951
	(32.4)	(31.9)		(31.8)	(31.0)	
n	845	836		1,704	1,711	

Notes: The table shows descriptive statistics of parental characteristics, which are not included in the estimation, at the cut-off for the estimation sample. It reports the variables' means before and after the school entry cut-off and the standard deviations are in parentheses. Columns 3 and 6 report the p-values of simple t-tests for differences in the variable' means before and after the school entry cut-off. Note that sample sizes vary by outcome. The number of individuals in the last row and variable statistics are based on the sample for the outcome of smoking behaviour.

Table A-5: Fuzzy RDD: including all observations per respondent in the estimation and sensitivity by age groups (four-month window).

	The sample includes							
	All ages		Age $\geq$ 30		Age $\geq$ 40		Age $<$ 60	
	1st obs.	All obs.	1st obs.	All obs.	1st obs.	All obs.	1st obs.	All obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Smoking behavior							
School starting age	-0.009^{**}	-0.009^{**}	-0.007	-0.009^{*}	-0.008	-0.009^{*}	-0.009^{**}	-0.010^{**}
	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)
1st stage: F-Statistic	246.2	146.0	151.1	97.6	119.8	81.8	238.7	144.1
n	3,449	10,968	2,208	7,350	1,538	5,247	3,315	10,630
	Self-rated health							
School starting age	0.025^{***}	0.019^{***}	0.019^{**}	0.016^{**}	0.027^{***}	0.015^{**}	0.021^{***}	0.015^{**}
	(0.007)	(0.007)	(0.009)	(0.007)	(0.010)	(0.007)	(0.007)	(0.007)
1st stage: F-Statistic	284.1	160.0	164.4	140.4	123.2	141.1	275.8	145.6
n	3,856	22,260	2,396	18,631	1,657	16,527	3,714	10,654
	Self-rated health: at least good							
School starting age	0.009^{***}	0.008^{**}	0.006	0.007^{**}	0.008	0.007^{*}	0.008^{**}	0.007^{**}
	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
1st stage: F-Statistic	284.1	160.0	164.4	140.4	123.2	141.1	275.8	145.6
n	3,856	22,260	2,396	18,631	1,657	16,527	3,714	10,654
	SF12: physical health score							
School starting age	0.148^{**}	0.129^{*}	0.146	0.140^{*}	0.192^{*}	0.144	0.101	0.107
	(0.069)	(0.067)	(0.090)	(0.085)	(0.104)	(0.098)	(0.069)	(0.067)
1st stage: F-Statistic	242.0	153.2	153.0	103.0	121.1	83.0	233.6	151.2
n	3,391	9,438	2,215	6,577	1,561	4,787	3,258	9,107
	SF12: mental health score							
School starting age	0.007	0.024	-0.045	-0.055	-0.077	-0.077	-0.014	0.026
	(0.081)	(0.075)	(0.097)	(0.087)	(0.107)	(0.097)	(0.083)	(0.076)
1st stage: F-Statistic	242.0	153.2	153.0	103.0	121.1	83.0	233.6	151.2
n	3,391	9,438	2,215	6,577	1,561	4,787	3,258	9,107
Birth year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey year indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Federal state indicators	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Migration background	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education father	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Education mother	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the second stage results of the fuzzy regression discontinuity design restricting the sample to the four-month window around the school entry cut-off. Analogous to the main analysis, '1st obs.' columns include only the first observation per respondent and 'All obs.' columns include all available observations per respondent in the estimation. Robust standard errors are in parentheses. The models including all waves use clustered standard errors at the level of respondents. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A-6: Fuzzy RDD: network of friends (all months and trends).

	No trend	General linear trend		Separate linear trends	
	(1)	(2)	(3)	(4)	(5)
			Number of friends		
School starting age	-0.082	-0.067	-0.069	-0.064	-0.067
	(0.052)	(0.066)	(0.064)	(0.064)	(0.062)
1st stage: F-Statistic	288.6	217	240.9	194.2	214.7
n	9,595	9,595	9,595	9,595	9,595
			Average age of friends		
School starting age	-0.184	-0.340^{**}	-0.182^{**}	-0.326^{**}	-0.168^{**}
	(0.145)	(0.156)	(0.075)	(0.164)	(0.078)
1st stage: F-Statistic	146.9	119.9	127.3	108.5	114.7
n	4,659	4,659	4,659	4,659	4,659
			Relative age of friends		
School starting age	-0.000	-0.004^{*}	-0.005^{**}	-0.004	-0.004^{*}
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
1st stage: F-Statistic	146.9	119.9	127.3	108.5	114.7
n	4,659	4,659	4,659	4,659	4,659
Birth year indicators	No	No	Yes	No	Yes
Survey year indicators	No	No	Yes	No	Yes
Federal state indicators	No	No	Yes	No	Yes
Female	No	No	Yes	No	Yes
Migration background	No	No	Yes	No	Yes
Education father	No	No	Yes	No	Yes
Education mother	No	No	Yes	No	Yes

Notes: Family members and relatives are excluded. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

hche | Hamburg Center for Health Economics

hche Research Paper Series, ISSN 2191-6233 (Print), ISSN 2192-2519 (Internet)

2011/1 Mathias Kifmann and Kerstin Roeder, Premium Subsidies and Social Insurance: Substitutes or Complements? March 2011

2011/2 Oliver Tiemann and Jonas Schreyögg, Changes in Hospital Efficiency after Privatization, June 2011

2011/3 Kathrin Roll, Tom Stargardt and Jonas Schreyögg, Effect of Type of Insurance and Income on Waiting Time for Outpatient Care, July 2011

2012/4 Tom Stargardt, Jonas Schreyögg and Ivan Kondofersky, Measuring the Relationship between Costs and Outcomes: the Example of Acute Myocardial Infarction in German Hospitals, August 2012

2012/5 Vera Hinz, Florian Drevs, Jürgen Wehner, Electronic Word of Mouth about Medical Services, September 2012

2013/6 Mathias Kifmann, Martin Nell, Fairer Systemwettbewerb zwischen gesetzlicher und privater Krankenversicherung, July 2013

2013/7 Mareike Heimeshoff, Jonas Schreyögg, Estimation of a physician practise cost function, August 2013

2014/8 Mathias Kifmann, Luigi Siciliani, Average-cost Pricing and Dynamic Selection Incentives in the Hospital Sector, October 2014

2015/9 Ricarda Milstein, Jonas Schreyögg, A review of pay-for-performance programs in the inpatient sector in OECD countries, December 2015

2016/10 Florian Bleibler, Hans-Helmut König, Cost-effectiveness of intravenous 5 mg zoledronic acid to prevent subsequent clinical fractures in postmenopausal women after hip fracture: a model-based analysis, January 2016

2016/11 Yauheniya Varabyova, Rudolf Blankart, Jonas Schreyögg, Using Nonparametric Conditional Approach to Integrate Quality into Efficiency Analysis: Empirical Evidence from Cardiology Departments, May 2016

2016/12 Christine Blome Ph.D., Prof. Dr. Matthias Augustin, Measuring change in subjective well-being: Methods to quantify recall bias and recalibration response shift

2016/13 Michael Bahrs, Mathias Schumann, Unlucky to be Young? The Long-Term Effects of School Starting Age on Smoking Behaviour and Health

The Hamburg Center for Health Economics is a joint center of Universität Hamburg and the University Medical Center Hamburg-Eppendorf (UKE).

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

UKE HAMBURG

hche Hamburg Center for Health Economics

Esplanade 36
20354 Hamburg
Germany
Tel: +49 (0) 42838-9515/9516
Fax: +49 (0) 42838-8043
Email: info@hche.de
http://www.hche.de
ISSN 2191-6233 (Print)
ISSN 2192-2519 (Internet)

HCHE Research Papers are indexed in RePEc and SSRN.
Papers can be downloaded free of charge from http://www.hche.de.