Causal Mechanisms of Relative Age Effects on Adolescent Risky Behaviours

Luca Fumarco, Lukáš Laffers, and Francesco Principe*

Abstract. Age differences between classmates are attracting growing attention in academic research and public policy, yet their underlying mechanisms remain understudied. We examine how relative age affects adolescents' risky behaviors across Europe. Using Health Behaviour in School-Aged Children (HBSC) survey data and a two-stage least squares (2SLS) strategy, we provide causal estimates that isolate relative age effects from absolute age and season-of-birth confounders. Relatively younger students are significantly more likely to engage in behaviors such as substance use. To explore mechanisms, we apply two approaches. First, a novel use of causal mediation analysis *à la* Dippel et al. (2022) shows that academic self-concept, well-being, self-esteem, and peer support serve as amplifying channels for these effects. Differently, effects on sexual and aggressive behaviors seems to be mostly driven by maturity differences. Second, reduced-form analyses using data from a connected survey (i.e., the European School Survey Project on Alcohol and Other Drugs (ESPAD)) suggest that younger students perceive lower risks and higher peer prevalence of substance use. These findings reveal psychological and perceptual pathways through which relative age influences adolescent behavior.

Keywords: Risky behaviors; Causal mediation analysis; Relative age effects; Enrollment cutoffs.
JEL Codes: C26; I10; I12; I21.

*Fumarco: Masaryk University, GLO, IZA, J-PAL, Lipová 41a, 602 00 Brno-střed, luca.fumarco@econ.muni.cz.
Laffers: Matej Bel University, Tajovského 40, Banská Bystrica, Slovakia and Masaryk University, lukas.laffers@gmail.com
Principe: University of Bergamo, Via dei Caniana, 2, 24127 Bergamo BG, Italy, francesco.principe@unibg.it.

Acknowledgements: We are grateful to Patrick Button, Ingo Isphording, Giacomo Pasini, Amelie Wuppermann alongside participants at the AIES 2021 conference, HERB workshop 2023, 8th LEER conference, EALE 2023 conference, and at the seminars at the JRC Ispra, University of Bergamo, EuHEA Seminar series for helpful suggestions and feedback. Luca Fumarco acknowledges the generous support from the NPO *"Systemic Risk Institute"* number LX22NPO5101, funded by European Union - Next Generation EU (Ministry of Education, Youth and Sports, NPO: EXCELES) and from the *CERGE-EI Foundation Teaching Fellowship Program*. Lukáš Laffers acknowledges support provided by the *Slovak Research and Development Agency* under contract no. APVV-21-0360 and VEGA-1/0398/23. Luca Fumarco and Lukáš Laffers acknowledge that this work was supported by the ERDF/ESF project *Ageing of the population and related challenges for health and social systems* (AGEING-CZ) (No. CZ.02.01.01/00/23_025/0008743). We acknowledge the work of the *HBSC study staff*, who conduct the survey in collaboration with WHO/Europe. The International Coordinator is Candace Currie, and the Data Bank Manager is Oddrun Samdal. A complete list of study coordinators, data bank managers, researchers, and participating countries is available on the HBSC website. We acknowledge the *ESPAD study staff*; in particular, Sabrina Molinaro (ESPAD coordinator), João Matias (EMCDDA representative), Sharon Arpa, Pavla Chomynova, Ludwig Kraus, Ola Ekholm, Julian Strizek, and Loredana Fortunato. This paper uses data from SHARE (DOIs: 10.6103/SHARE.w2.900, 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800, 10.6103/SHARE.w8.800, 10.6103/SHARE.w9.800). The SHARE data collection has been primarily funded by the European Commission through FP5, FP6, FP7, Horizon 2020, and Horizon Europe. Additional funding from national sources is gratefully acknowledged (see www.share-project.org. For methodological details, see Börsch-Supan et al. (2013).

1 Introduction

There is a vibrant public debate on within-class age differences, especially in the US, which has reinvigorated in the aftermath of the COVID-19 pandemic.[1] These within-class age differences (henceforth referred to as relative age) are defined by students' ages compared to those of their classmates. Research has shown that relative age can have far-reaching effects on individuals' educational performance (Bedard & Dhuey, 2006; Fredriksson & Öckert, 2014), as well as their mental and physical health (Schwandt & Wuppermann, 2016; Arnold & Depew, 2018). This study contributes to the existing literature by investigating the effects of relative age on risky behaviors and, more importantly, exploring possible mechanisms.

Studies have shown that relative age is one of the leading cause of risky behaviors (Johansen, 2021; Argys & Rees, 2008). However, the literature on this topic has at least one limitation: the mechanisms behind these relative age effects on risky behaviors are still speculative. The limited empirically-driven discussion with respect to mechanisms limits policy makers' ability to design appropriate interventions.[2] The aim of this paper is to fill this gap in the literature.

We use data from the international survey *Health Behaviour in School-Aged Children* (HBSC) on adolescents from more than thirty European countries. We start by employing the two-stage least squares (2SLS) methodology to address the endogeneity of relative age and find a series of robust results. Relatively young students are more likely to smoke tobacco and drink alcohol. This is the first paper to provide a causal estimate of relative age effects on marijuana use, showing that relatively young students are more likely to consume this substance. Unlike previous literature, our results suggest a less risky sexual life for relatively young students in terms of condom use, on average in Europe. Finally, relatively old students are more likely to bully and to be involved in fights, although they do not differ from younger classmates in terms of cyberbullying. These main analyses contribute to the literature by addressing both the endogeneity of relative age and isolating its effect from that of absolute age.

In the second part of the study, we investigate mechanisms in two ways. First, we contribute to the small but growing applied economics literature that uses causal mediation analyses, with Dippel et al. (2022) being one of the most prominent studies in applied economics. In our analyses, we explore the mediating effects of academic self-concept, well-being, self-esteem, autonomy with respect to risky health behaviors, and peer support.

This investigation on the role of the mediators returns policy relevant results. We find that students' academic self-concept, well-being, self-esteem, and peers' support are important indirect channel through which relative age affects risky health behaviors, in terms of substances consumption. This finding implies that simple solutions could be adopted to mitigate risky behaviors in adolescence.

We believe that the results could be causally interpreted. In our analysis, we follow Dippel et al. (2022)'s framework, which allows the usage of a single instrument for both the mediator and the endogenous treatment. This is due to the specific structure that is imposed: there should be no unobservable variable that is orthogonal to the mediator and

[1] It is discussed in various news outlets, such as USNews and The Guardian, and at any governmental level, from school districts to intragovernmental institutions, such as UNESCO.

[2] A rare exception is the study from Page et al. (2017), who conduct experiments to study relative age effects on risky attitudes, and risk-taking attitudes.

that affects both the outcome and the treatment (in our case, relative age), conditional on the set of control variables. In the text, we argue why we think this is a plausible assumption.

Second, we further study mechanisms behind relative age effects on consumption of addictive substances. In particular, this is the first study to investigate whether relative age is associated with perceived risk of consumption and perceived peers' prevalence of consumption. For these reduced-form analyses, we use comparable survey data from the *European School Survey Project on Alcohol and Other Drugs (ESPAD)* and find two results. First, relatively young students perceive lower risks associated with the consumption of addictive substances. Second, compared to their older classmates, they perceive that the prevalence of consumption of these substances is higher among their peers. The combination of these results suggests that the effect of relative age on addictive substances consumption could be countered with informational campaigns, specifically targeting the youngest students in a cohort.

By investigating relative age effects on adolescents' risky behaviors and its mechanisms, we also contribute to the literature on the role of the educational system and peers in determining individual risky behaviors. For example, Card & Giuliano (2013) find significant peer effects in adolescents' sexual initiation, smoking, marijuana use, and truancy. These effects are greater for females. Elsner & Isphording (2018) find a negative effect of high school students' rank on the likelihood of smoking, drinking, having unprotected sex, and engaging in physical fights. Reynoso & Rossi (2019) exploit a natural experiment in Buenos Aires and show a relationship between attending high school at night and the probability of having unsafe sex or consuming substances, finding evidence on the absence of parental supervision as a key mechanism.

The literature on the effects of relative age on risky behaviors is characterized by two additional limitations that we address in this paper. First, relative age is usually not disentangled from other age effects. So far, only a few studies have conducted this disentanglement (e.g., Peña & Duckworth (2018); Peña (2017); Cascio & Schanzenbach (2016); Black et al. (2011)). However, this disentanglement is policy relevant (Dhuey, 2016). For example, while it is possible to mitigate an individual student's age at school start effect by postponing her school entry (i.e., redshirting), this postponement would have an ambiguous impact on her classmates, in terms of relative age effects.[3] Because of the particular features of the HBSC survey, we can study relative age in isolation.

Second, since HBSC smallest sample unit is the class, we can interpret our results in terms of proper peer effects compared to the analysis of students within a grade or cohort through classes, who might never interact in real life. There is broad consensus among scholars that classroom interactions are important determinants of risky health behaviors (Balestra et al., 2021).

The remainder of the paper is organized as follows. In the next section, we discuss the related literature on age at school start, relative age, and determinant of risky behaviors.

[3]In other words, while redshirting might seem like a way to reduce an individual student's relative age disadvantage, it doesn't eliminate the relative age effect. Instead, it shifts the advantage from one student to another, creating a zero-sum game. For example, imagine Student X, who was born 30 days before the school cutoff and is redshirted. In her new class, she becomes the oldest student, 30 days older than Student Y, who was born exactly on the cutoff date and would have been the oldest in a typical class. Meanwhile, Student Z, who was born one day before the cutoff, was already nearly a year younger than Y. Now, Z is 394 days younger than X. This shift shows how redshirting benefits the individual (X) but widens the age gap for others (like Z), potentially worsening the relative age effect for the class as a whole.

Section 3 describes the data, whereas Section 4 presents the empirical strategy. Section 5 presents main results, while Section 6 discusses potential mechanisms. The final section provides a summary and concludes.

2 Literature Review

In this section, we discuss in greater detail the differences between this study and the small literature on relative age effects on risky behaviors. Johansen (2021) uses Danish register data and a fuzzy regression discontinuity design and find that women who were the youngest children in their cohort are more likely to abort and experience alcohol poisoning. Argys & Rees (2008) use data from the National Longitudinal Survey of Youth and find that female students with younger peers use substances more frequently than female students with older peers, while there is no equivalent effect on male students.

Argys & Rees (2008) uses a reduced form analysis; however, differently from Johansen (2021), it disentangles relative from absolute age effects. In the literature, some terms referring to age effects tend to be used interchangeably, although they do not represent exactly the same thing. Relative age in isolation is a peer effect; it is an externality generated by being older than classmates. While the relative age effect is the effect of age differences between classmates, there are other three age effects in the literature. First, age-at-school start effect, that is, the effect of the age at which students start school, and this is what the literature usually focuses on, as in Johansen (2021). Typical studies that use a regression-discontinuity design (RDD) to investigate age at school start provide an estimate which includes relative age: two students born around the cutoff have about a similar age-at-outcome when they are observed and they are born almost at the same time, but they started school one year apart (one of them might also have spent one additional year in the kindergarten) *and* they are on the opposite relative age spectrum within their school class. Second, age-at-outcome effect, that is, the proper absolute age effect: the effect of the age at which the outcome was measured (or the survey was conducted). Third, time-in-school effect, that is, the effect of the time spent in school. Typically, these different but related age effects cannot be all disentangled at the same time; a rare exception is Nicodemo et al. (2024), on age effects on ADHD diagnosis.

There is one additional and recent study on risky behaviors, that is, Lopez-Mayan et al. (2024). The authors investigate age-at-school start effects and uses an RDD as in Johansen (2021); thus, similarly to Johansen (2021), it differs from our study in methodological terms and in the interpretation of the estimates. This manuscript has a number of novel aspects; most importantly, it investigates previously unexplored outcomes (e.g., gambling, gaming, vaping) and focuses on how the estimates of age-at-school start effects are influenced by institutional aspects, such as progressing in the educational system and retention. Lopez-Mayan et al. (2024) finds that younger for grade students are less likely to engage in risky behaviors, and that these gaps vary with students' grade and absolute age.

Due to the features of the dataset at hand, we are able to isolate relative age from the other three age effects. In our setting, countrywide (expected) age at school start–which is given by country-specific regulations–is captured by country fixed-effects. Moreover, our estimate of absolute age effects combines age-at-outcome and time-in-school–conditional on country's school entry requirements; thus, we cannot measure these two effects separately, but this is not a problem because they are not the focus of our paper. With this respect, our setting is closer to Argys & Rees (2008) than to Johansen (2021) and Lopez-Mayan et al. (2024).

There are four more studies that focus on a subset of risky health behaviors. Routon & Walker (2022) find mild evidence that the old students in a grade drink more alcohol. Ballatore et al. (2020) and Mühlenweg (2010) find that the young students in their school grade are more likely to be victimized. Bahrs & Schumann (2020) find that adults in their mid-thirties, who were the old students in their school grade, are less likely to smoke. While these studies address the endogeneity of relative age, do not effectively disentangle relative age from absolute age and investigate a limited set of risky health behaviors.

Finally, most of the above studies do not effectively investigate mechanisms through which relative age affects risky behaviors. Page et al. (2017) is an exception with this respect. They conduct lab experiments in Australia and find two types of results: (i) negative relative age effects on risky attitudes, concerning events out of people's control (i.e., BART test, where participants guess the number of pumps to a balloon before it explodes), and (ii) positive effects on risk-seeking attitudes, where the perceived level of risk depends upon respondents' experience (e.g., risky behaviors while driving), which might be given by greater self-confidence.[4]

Compared to the above studies, we address endogeneity of relative age and separate its effect from that of various absolute age concepts; moreover, we study potential mechanisms.

3 Data and Variables

We use data from the *Health Behaviour in School-Aged Children (HBSC)*, a cross-national survey on adolescents' health and well-being, administered in schools every four years. In our analysis, we use five consecutive waves: 2001/2, 2005/6, 2009/10, 2013/14, and 2017/18. We exclude countries for which we have unbiguous information on the school enrollment cutoff date (e.g., Albania, Arzebaijan, Serbia), for which the cutoff date falls in the middle of the month–because we do not have the precise day of birth (e.g., Portugal, Roumania, Israel),[5] or that adopts multiple cutoffs within country (e.g., Switzerland, Germany, US, Canada)–because we do not have information on the region or state of students' schools. The final sample is composed of more than 600,000 students from 30 European countries with vastly different characteristics, which qualifies our study as a multi-country study. This is a rarity in the literature on both relative age and age-at-school start. Table B.1 in the Appendix provides the number of observations by country and by wave, with country-specific cutoff dates.

The sampling unit at the class level, and the large variation in both absolute age and country-specific cutoff dates, makes the HBSC data particularly attractive. Respondents' age ranges between 10.5 and 17.5 years, and cutoff dates–as well as age at school entry–vary across countries. These sources of variation allow us to disentangle relative age from both absolute age and season-of-birth effects, as discussed below in greater detail.

[4] This study refers to relative age effects; however, it exploits the school enrollment cutoff date, comparing students born just before and after it. This approach means that students in different grades are being compared (e.g., the oldest in grade 8 versus the youngest in grade 9). Thus, results from this study are more comparable to those on age-at-school start effects on risky behaviors, such as Johansen (2021) and Lopez-Mayan et al. (2024), than to our results and those in Argys & Rees (2008).

[5] This missing information precludes us from using the RDD and to study age at school entry effects, and then compare them to relative age effects in isolation.

Outcome variables. We study the probability of having conducted various risky behaviors: *(i)* Early smoking, a dummy variable which equals one if the student has already smoked once by 13 years of age; *(ii)* Smoking currently, a dummy variable which equals one if the student currently smokes; *(iii)* Early drinking, a dummy variable which equals one if the student has already drunk once by 13 years of age; *(iv)* Ever drunk, a dummy variable which equals one if the student has been drunk at least once in life; *(v)* Early Marijuana, a dummy variable which equals one if the student has smoked marijuana at least once by 13 years of age;[6] *(vi)* Ever smoked marijuana, a dummy variable which equals one if the student has smoked marijuana at least once in life; *(vii)* Ever sex, a dummy variable which equals one if the student has had sex at least once in life; *(iix)* Unprotected sex, a dummy variable which equals one if the student has had sex without a condom during the last intercourse; *(ix)* Fight, a dummy variable which equals one if the student was involved in at least one fight in the past year; *(x)* Bully, a dummy variable which equals one if the student has bullied someone at least once during the past two months; *(xi)* Been bullied, a dummy variable which equals one if the student has been bullied at least once during the past two months; *(xii)* Cyberbullied, a dummy variable which equals one if the student has been cyberbullied (either with pictures or messages) at least once in life.

Sample sizes varies across analyses of these outcomes. Questions on *(i)* to *(iix)* are asked only in classes where expected students are at least 15 years old at the moment of the survey. Questions on *(i)*, *(ii)*, *(iii)*, and *(iv)* are not asked in wave 2018. Outcome *(v)* is based on observations from wave 2014 only. Outcome *(xii)* is based on observations from waves 2014 and 2018 only. For these reasons, our analyses always study a sample that is smaller than the initial raw sample size, that is, 600,000 students.

To the best of our knowledge, relative age effects on outcomes *(ix)* and *(xii)* have not been previously studied.[7] However, they provide useful insights. Participation in fights is a more precise indicator of physical violence than bullying; this is important, because physical violence between underage people can lead to legal consequences in some places.[8] Cyberbulling is on the rise since the increase in social media platform usage.[9]

Table 1 reports the number of observations and means per risky health behavior.

*****Table 1 about here*****

Relative age. Relative age, RA_{ic}, is measured as the difference between the age of the student i's in class c, AGE_{ic}, and the age of the oldest regular student j in class c, AGE_{jc}; thus, this measure varied by class and its increase implies that student i is relatively older. By "regular student" we mean that the student is in the right class based on their age and on the country cutoff date. Note that this measure of relative age, which

[6] Three outcomes focus on 13 years of age as threshold based on findings from the medical literature. For example, the HBSC created a variable on smoking before the age of 13 because this is the peak age for smoking experimentation (Duncan et al., 1995), and because there is evidence that early initiation to tobacco use seems to be linked to a higher likelihood of persistent smoking (Ellickson et al., 2001).

[7] Similarly, we are not aware on any study on age at school start on these two outcomes.

[8] For example, in European countries and in some US states, such as in Illinois, students can be accused of assault and battery.

[9] This was particularly true during the worse time of the COVID-19 pandemic, as discussed by the UN; however, this most recent period is not covered by our data.

varies by class, can be constructed because the primary sample unit in the HBSC survey is the class.[10] More formally, relative age is constructed as in Equation 1:

$$RA_{ic} = AGE_{ic} - \max\{AGE_{jc} : j \in R_c\}. \tag{1}$$

For regular students i in class c, $i \in R_c$, this measure ranges between zero (i.e., student $i = j$, she is the oldest regular student in the class) and -1 (i.e., there is almost one year difference between student i and the oldest regular student in the class, student j).[11]

There are several reasons why relative age, RA_i, is likely to be endogenous. Parents may greenshirt (i.e., expedite school entry) their children or have them skip a grade. At the opposite, parents may redshirt (i.e., delay school entry) their children or teachers may decide to retain them. On one hand, children who begin school early or skip a grade are likely to be born right after the cutoff date, and thus to be perceived as being especially skilled compared to (younger) children in the same school cohort; on the other hand, children who begin school later or are retained are likely to be born right before the cutoff date, and thus to be perceived as having developmental problems with a higher frequency than (older) children in the same school cohort (Peña, 2017; Schwandt & Wuppermann, 2016; Sprietsma, 2010; Bedard & Dhuey, 2006). On top of that, in some countries the decision to redshirt may vary by parents' socio-economic status (SES) (Bedard & Dhuey, 2006). We address concerns of endogenous selection into treatment by implementing an instrumental variable strategy, where the instrument is expected relative age.

Table 2 reports observations, means, and standard deviations for relative age and the other independent variables.

*****Table 2 about here*****

Regular students represent 83% of the sample, while non-regular students represent about 17% of the students. The percentage of non-regular students with extreme values is very low,[12] and their inclusion in the analyses does not alter the results.

As a final note, this construction of relative age exploits within-classes variation. Therefore, we could be able to separate relative age from absolute age even if we did not have cutoff date variation (i.e., students in the same country and grade, with the same absolute age, who entered school in the same year, could have different relative age, because of the different class-specific age composition).

[10] The class size ranges between a minimum of 8 and a maximum of 31 students, with median and mean of 18 students, with a standard deviation of about 5.

[11] It is *almost* one year, because exactly one year would mean that student i was born on the same day, but in the next academic year.

[12] We define extreme relative age values as those below -2 or above 1. Thus, this definition of extreme values of relative age pertains neither students with relative ages between 0 and 1 nor students between -1 and -2. While students with relative ages between 0 and 1 do not fall within the expected range, they are typically retained or redshirted by one year. Similarly, students with relative ages between -1 and -2 are either greenshirted or have skipped one grade. Students with extreme relative age values are most often older–rather than younger–than expected and may have been retained multiple times and/or face additional learning disadvantages beyond those associated with lower relative age (e.g., ADHD). These students with extreme values of relative age make up 1.6% of the sample. Among them, 0.2% are younger than expected, while 1.4% are older than expected. When considering only students with a relative age above 2 (i.e., at least two years older than expected), this proportion drops to 0.2%.

Expected relative age. Expected relative age, ERA_{iCOU}, is the month of birth of student i within the academic year of country COU and measures the number of integer months between the cutoff date and the academic month of birth of student i, as if this student entered school when they were supposed to.[13] This discrete variable ranges between zero and 11, with zero being the reference month that starts with the cutoff date and 11 the month that precedes the cutoff date. Therefore, ERA does not vary by class, but it varies by country due to the country variation in cutoff dates.

Our main analyses use a disaggregated version of this variable, that is, a vector of dummies for the academic month. This is suggested in Angrist & Pischke (2008), and it is equivalent to Angrist & Krueger (1991), who disaggregated the calendar year into quarter of birth dummies and then conducted overidentification tests. This categorization is not arbitrary: it is month of birth normalized based on the country-specific cutoff date.[14]

Let us make a practical example to see how this instrument is constructed. Consider a student born in December in a country with September 1st as the cutoff date (e.g., Luxembourg). Their ERA would be three, that is, she was born three months after the month starting with the cutoff date, for which ERA is set to zero. In the disaggregated version of this variable, the dummy for the academic month of birth three equals one, while dummies for other academic months of birth equal zero.

The disaggregation of expected relative age provides us with two main benefits. First, this transformation increase the first-stage fit. Second, it allows us to test the validity of the instruments with a standard overidentifying restriction test; this is because there are more instruments than endogenous variables.

Absolute age. Absolute age, AA_i, captures a mix of age at the time of survey-participation and time spent at school–conditional on country school entry rules, which–similarly to RA–might depend on parents SES.[15,16] In any given class, students from high SES families might tend to be older (younger), if these families redshirt (greenshirt, depending on the country) their children more frequently. Thus, we instrument this variable with expected absolute age. We conduct additional analyses treating absolute age as exogenous; these analyses return virtually identical results, and serve to fit the

[13]This instrument, or similar versions, have already been used in the literature in Fumarco et al. (2020); Fumarco & Baert (2019); Page et al. (2019); Peña & Duckworth (2018); Peña (2017, 2020); Datar (2006). The latter studies measure ERA as the distance in non-integer years between student i's age–if they were a regular student–and the age of the hypothetically youngest (oldest) student in the class, who was born right before (right on) the cutoff date. This variable can be used also as variable of interest in reduced form studies, such as Cygan-Rehm & Westphal (2024) and Barabasch et al. (2024).

[14]Similar fixed effects instruments are used in the so-called "many-weak-instruments" literature, where judge or examiner fixed-effects are directly used as instruments or to create instruments (Kling, 2006; Maestas et al., 2013; Aizer & Doyle Jr, 2015; Dobbie et al., 2018; Agan et al., 2023; Mueller-Smith, 2015; Bhuller et al., 2020; Di Tella & Schargrodsky, 2013). Also part of Angrist & Krueger (1991) could be put into this literature; they use various transformations of quarter of birth instruments and, with the increases in the number of instruments–and the decrease in observations per instrument–the many-weak-instrument inference problem surfaces (e.g., by interacting quarter with year of birth and place of birth) (Angrist & Frandsen, 2022). Differently from this literature, our dummies for academic month of birth are not weak instruments, as shown in Table B.2 and Tables B.3, B.4, and B.5.

[15]We explicitly focus on relative age effects, so the fact that we cannot separate the effects of age at the time of survey-participation and time spent at school does not matter for the sake of this study.

[16]Most other studies using this dataset include absolute age as a control variable (Fumarco et al., 2020; Fumarco & Schultze, 2020; Fumarco & Baert, 2019; Fumarco et al., 2024; Carpenter & Churchill, 2024). Since we are using the same dataset, we also center absolute age. In the previous studies, we did this to explore heterogeneous effects and include interaction terms. However, using non-centered age instead would not meaningfully affect our results.

mediation analyses within Dippel et al. (2022)'s causal framework.[17]

Expected absolute age. Expected absolute age EAA_i represents the absolute age that student i would have had in that class, if she was a regular student. It is the expected absolute age of students who participate in the same survey, live in the same country, attend the same classroom, and were born in the same quarter; thus, it is based on classmates age. This specification closely follows Peña & Duckworth (2018).

Control variables. We control for the standard set of demographic characteristics. First, we control for students' gender. This variable equals one for female students and zero for male students, the reference group.

Second, we control for whether the student lives with both parents. Note that a recent Danish study find that relative age may affect marriage stability (Landersø et al., 2020); however, balance tests suggest that relative age and family status are not related in our multi-country sample, see Section A.2. Moreover, we conduct robustness checks where we exclude this control variable and the main results remained unchanged, as expected by Frisch–Waugh–Lovell theorem.

Third, we account for the SES of students' families. SES is derived from multiple items according to the HBSC guidelines (C. Currie et al., 2008) and it is coded into three dummies for high, medium, and low SES, with the latter being the reference category. Note that a family's SES might be endogenous: it is the family's SES when the survey was conducted and, potentially, relative age might have influenced the family's SES since birth, through the influence on the mother's labour market outcomes (Landersø et al., 2020). We address this concern in a robustness check, where we exclude family SES from the control variables. These results are indistinguishable from the main results. Related to the above, SES might be suspected to affect RA and, more importantly, ERA: some studies find that families with certain SESs might target different dates of delivery. To address this concern, we conduct balance tests on ERA; these results suggest that in Europe, on average, families with a different SES do not target different dates of birth. The related literature and these tests are discussed in greater detail in Section A.2. Possible differences between countries are captured by country fixed effects.

Analyses also account for unobservable birthdate effects, known as "season-of-birth effects." The variable for season of birth is proxied by the month of birth within the calendar year (henceforth calendar month) and ranges between zero (January, the reference month) and 11 (December). If left unaccounted for, season-of-birth effects could cause biased estimates. For instance, Bound & Jaeger (2000) explain that individuals born in wintertime more likely suffer from multiple health issues, such as mental disabilities and multiple sclerosis, while individuals born in Spring are more likely to be shy. Thus, season-of-birth effects capture calendar-period specific effects on health outcomes that do not depend on maturity differences, but that might cause differences between students born in different periods of the year. Other studies providing evidence of season-of-birth effects are J. Currie & Schwandt (2013), which investigates the effect of time of conception on birth-weight, and Dustmann et al. (2022), Trudeau et al. (2016), and Wernerfelt et al. (2017), which study the effect of sunlight during pregnancy and infancy on asthma, birth-weight, and childhood obesity.

It is important to highlight that since our empirical setting leverages variation in cutoff dates, expected relative age does not overlap with the calendar month of birth. In Figure A.1, Section A.1, in the Appendix illustrates some examples.

Finally, the analyses account for wave and country fixed effects. Among other things,

[17] In any country, grade would be collinear with absolute age, so we cannot control for grade as well.

country fixed effects capture cross-country variation in statutory (integer) age at school entry and for country-specific law and social norms that might affect risky behaviors. Wave fixed effects capture possible time specific shifts in risky behaviors.

Mediator variables. We investigate the mediation effect of five channels: *(a)* student's academic self-concept in relation to the classmates; *(b)* perceived well-being; *(c)* satisfaction with own body image; *(d)* time spent outside with friends in the evening; and *(e)* perceived support from other students.[18] Channel *(a)* is a measure of academic self-concept. It is given by the standardized output of the survey question asking students what their teachers might think about their school performance compared to classmates'. The original variable ranges from zero (below average) to three (very good). Channel *(b)* is a proxy for well-being. It is an index created by applying the principal component analysis to the standardized output of survey questions asking students about their perceived life-satisfaction and self-rated health, which are closely associated variables. The original variable of life-satisfaction ranges from zero (worst possible life) to ten (best possible life). The original variable of self-rated health ranges from zero (poor) to three (excellent). Channel *(c)* is an index that measures self-image. It equals one if the student thinks their body is about the right size and zero otherwise (i.e., much or a bit too thin, much or a bit too fat). Channel *(d)* is a proxy for parents' supervision. It is the standardized output of the survey question asking students how many days per week they meet with friends after 8pm. The original variable ranges from zero to seven days per week. Channel *(e)* is a measure of peers' acceptance. It is an index created by applying the principal component analysis to the output of survey questions on schoolmates enjoying spending time together, helping each other and being kind, and accepting other students as they are. The three original variables ranged from zero (strongly agree) to four (strongly disagree).

4 Empirical Strategy

To study the impact of a student's relative age on risky behaviors, we estimate relative age effects with a 2SLS.

Had we to use a OLS, the regression specification would be the one below:

$$Y_i \quad = \quad \beta_Y^0 + \beta_Y^{RA} RA_i + \beta_Y^{AA} AA_i + \beta_{\mathbf{Y}}^{\mathbf{X}} \mathbf{X_i} + \epsilon_i^Y \tag{2}$$

Index i indicates the individual and Y_i is one of the outcome variables measuring risky behaviors discussed in Section 3. We use RA_i and AA_i to denote the measures

[18] This list of mediators is not meant to be complete. We choose these mediators based on a survey of evidence from other fields, as they provide insights into various psychological and social factors that can contribute to the relationship between relative age and risky behaviors, helping to explain the underlying mechanisms at play. In fact, we chose these mediators because: (a) student's academic self-concept in relation to their classmates can influence risk-taking to assert status (Elsner & Isphording, 2018; Massey et al., 2008); (b) low levels of perceived well-being may lead to risk-seeking behaviors (Sorbring et al., 2014; Santini et al., 2020); (c) satisfaction with own body image can drive risky behaviors to conform to physical ideals (Granner et al., 2002; Wild et al., 2004; Woertman & Van den Brink, 2012; Gillen et al., 2006); (d) time spent outside with friends in the evening provides opportunities for risky behaviors due to increased independence and lack of parental supervision (Reynoso & Rossi, 2019; Averett et al., 2011; Fletcher, 2012); and (e) perceived support from other students acts as a protective factor against risky behaviors through a sense of belonging and reduced likelihood of engaging in detrimental activities (Springer et al., 2006).

of relative and absolute age, respectively. Vector of covariates $\mathbf{X_i}$ includes dummies for gender, living with both parents, family SES and also fixed effects for the calendar month of birth, survey wave, and country. Since relative age is measured at the class level, standard errors are clustered at the class level throughout.

Equation 2 would provide biased estimates of β_Y^{RA} and β_Y^{AA}, due to the endogeneity of relative and absolute age; thus, we resort to using a 2SLS approach. Equations 3 and 4 specify the first stages for the two endogenous variables, that are instrumented using a vector of dummies for expected relative age ($\mathbf{ERA_i}$) and expected absolute age (EAA_i).

$$RA_i = \beta_R^0 + \beta_R^{ERA}\mathbf{ERA_i} + \beta_R^{EAA}EAA_{iCOU} + \beta_R^X\mathbf{X_i} + \epsilon_i^R \tag{3}$$

$$AA_i = \beta_A^0 + \beta_A^{ERA}\mathbf{ERA_i} + \beta_A^{EAA}EAA_{iCOU} + \beta_A^X\mathbf{X_i} + \epsilon_i^A \tag{4}$$

The second stage is given in Equation 5.

$$Y_i = \beta_Y^0 + \beta_Y^{RA}\widehat{RA}_i + \beta_Y^{AA}\widehat{AA}_i + \mathbf{X_i}\beta_\mathbf{Y}^\mathbf{X} + \epsilon_i^Y. \tag{5}$$

where $\widehat{RA}_i$ and $\widehat{AA}_i$ refer to first stage predicted values.

Before proceeding with the analyses, we conduct balance tests to verify that ERA is orthogonal with respect to observable characteristics. We find that ERA is balanced with respect to all observable characteristics, and in particular with parents' SES. These results on the unbiased nature of ERA, and thus on birth date exogeneity, extend to EAA as well. Section A.2 in the Appendix discusses these tests and the results in greater details.

For comparison sake, one could wonder about how relative age effects obtained with a 2SLS, while controlling for absolute age, compare to age at school start effects obtained with the similar fuzzy RDD. However, the dataset does not have information on date of birth, which is a fundamental piece of information for RDD analyses.[19]

5 Main Results

Table B.2 in the Appendix shows the two first-stage regressions. In Column (1), we observe that the effect of each dummy variable for ERA on observed RA increases *almost* monotonically. These estimates are highly statistically significant, and suggest the disaggregation into dummies could be more suitable than using the not-disaggregated version for RA. In Column (2), we observe that the correlation between EAA and observed AA is particularly high; this result is due to a large number of regular students, that is, students who are in the class where they were supposed to be, net of redshirting, greenshirting, retention, or grade skipping.

For brevity sake, Table 3 reports only the second-stage estimates of relative age effects for all outcomes. Full statistics, with estimates on control variables and 2SLS ancillary tests, are reported in the Appendix in Tables B.3, B.4, B.5.

[19]The dataset does not even have information on the exact day of survey participation, which could allow to estimate the date of birth, coupled with age.

*****Table 3 about here*****

Relative age coefficients can be interpreted as percentage changes in the probability to adopt a certain risky health behavior following an increase in relative age by one year (i.e., the hypothetical maximum age gap between regular students–about 12 months): in other words, its increase implies that the student is relatively older.

The results reveal a consistent pattern across outcomes. Relatively young students are more likely to engage in smoking and drinking; an increase by one year in relative age decrease the chances of smoking before turning 13 by about 4.4%, and the chances of currently smoking by 2.4%. An increase by one year in relative age decrease the chances of having been drunk at least once before turning 13 by 6.9%, and of being drunk at least once in life by 5%. These results are statistically significant at the 1% level. Relative to baseline probabilities, these are economically significant effects; this is particularly true for having been drunk at least once in life, which is 20% of the average probability, see Table 1.

Results on smoking marijuana are consistent with those on smoking tobacco and drinking. An increase by one year in relative age decrease the chances of having smoked marijuana at least once before turning 13 by about 1.4%, and of having smoked marijuana at least once in life by 2.2%. Both these results are statistically significant at the 1% level. An increase by one year in relative age increase the chances to have had sex at least once in life by 2.2%, and the chances to have had sex without a condom during the last intercourse by 5.5%. These results are statistically significant at the 5% level. Relative to baseline probabilities, these are economically significant effects, with the largest magnitude being that for having had sex without a condom in the last intercourse, which is 16% of the average probability, see Table 1.

Finally, an increase by one year in relative age decrease the chances of having been bullied by 1.2%, while it increases the probability of having bullied at least once in the previous two months by 2.2%. An increase in relative age by one year does not affect the probability of having been cyberbullied at least once in life, while it increases the probability of being involved in a physical fight by about 4.1%.

We can assess the economic significant of these results by comparing them to SES estimated effects as well. In particular, the absolute value of the relative age effects on smoking and drinking habits, as well as on early consumption of marijuana and being involved in fights, ranges between 2 and 14 times the effect of being from a high-SES rather than from a low-SES (see Tables B.3 and B.4).

We conduct various robustness checks. We repeat the main analyses without disaggregating ERA (Table B.6); the caveat is that we cannot conduct the overidentification test without disaggregating ERA. We investigate further relative age specifications (Figure B.4).[20] We conduct additional analyses while omitting family SES (Table B.7) or the presence of both parents at home (Table B.8). We explore whether relative age effects vary when controlling for class-level characteristics, namely class size, share of female students, and share of students with high SES (Table B.9). We study whether relative age varies while controlling for the quantity siblings at home and whether at least one

[20]In all these versions, the instrument is always called ERA, but it is constructed slightly differently: (i) and (ii), the difference between own age and the average age in class, with or without student i's age; (iii) the age difference from the age of the youngest regular student in class; or (iv) the within-class percentile age rank.

of them is older (Table B.10). We repeat the analyses with clustered standard errors at country level (Table B.11), with school fixed effects in lieu of country fixed effects (Table B.12),[21] without students with extreme relative age values (Table B.13), without instrumenting for absolute age (Table B.14),[22] and with an alternative version of EAA, that is, the median of classmates' (discrete) absolute age (Table B.15). By and large, the results are virtually identical. Additionally, leave-one-out robustness checks show that there is no single country driving the results (Tables B.1, B.2, and B.3).

Details on diagnostics are reported in Table B.3, B.4, and B.5 in the Appendix. For all analyses, under-identification tests reject the null hypothesis that the instruments are not correlated with the endogenous variable, while weak-identification tests suggests they are not weakly correlated.[23] Moreover, in most cases, overidentification tests fail to reject the null hypothesis that the instruments are uncorrelated with the second-stage error term. The overidentification test rejects the null hypothesis for: having smoked and drunk before turning thirteen, having ever smoked marijuana, having had unprotected sex, and having been bullied in the recent past. Thus, results on these outcomes should be considered with caution.

These main results are mostly statistically significant at conventional levels and their magnitude is economically significant. How do they compare to the closest studies? Results on smoking and drinking are in line with those from Elsner & Isphording (2018), who investigate the effect of performance on risky behavior among American adolescents. Results on smoking are in line with those from Bahrs & Schumann (2020).

The results on unprotected sex and other risky behaviors are opposite those of Johansen (2021) and Lopez-Mayan et al. (2024). We should highlight that our study differs from those two with respect to three main aspects. First, this result comes from a representative sample of adolescent students coming from most European countries, while Johansen (2021) use administrative data from Denmark and Lopez-Mayan et al. (2024) use survey data from Spain. Country-specific social norms and drop out policies–based on grade versus age–could make a difference. For example, Johansen (2021) writes that policies may affect differences between genders on risky sexual behaviors in Denmark, and, we think, they might play a key role in determining differences between countries too. Note that country fixed effects capture such differences in policies and/or norms. Second, differently from Johansen (2021), in our analyses, boys and girls are considered together. In two additional analyses, we study boys and girls separately. There, we observe no substantial difference between genders, with a point estimate of 0.061 for boys and 0.054 for girls, and statistical significance at 10% for both of estimates.[24] These

[21] We do not conduct robustness checks using class fixed effects for one key reason. Incorporating class fixed effects would absorb much of the variation in relative age itself, since relative age is calculated relative to a reference classmate. Moreover, it is worth noting that school fixed effects account for potential between-school sorting–such as parents' non-random choice of school based on SES or other factors. This sorting can result in students attending schools where peers have similar socio-economic backgrounds, which may in turn influence the prevalence of risky behaviors within the school (Lopez-Mayan et al., 2024). As such, school fixed effects might be more appropriate than class fixed effects.

[22] The idea is that, although endogenous, the small magnitude of this endogeneity does not make a difference for sake of the relative age effect estimates. This idea is applied also in other contexts: students' gender is not completely random in Western countries either (e.g., in the US), mainly due to medical reasons, but it is *always* treated as exogenous.This is not usually discussed, but the implicit reasoning is that the extent of this endogeneity is so small that it does not affect the estimated effect of the variable of interest.

[23] For the latter, the F statistics are well beyond critical values suggested in Stock & Yogo (2005).

[24] While this paper does not include these results, the replication package includes the Stata do file for

results by gender should be considered with caution: the overidentification test reject the null hypothesis for females. Third, we disentangle relative from absolute age and we focus on within-class age differences. Finally, it should be noted that, although Argys & Rees (2008) use a reduced form analysis, they separate absolute and relative age as we do, and their results are equivalent to ours–with relatively younger students engaging more in risky behaviors.

Our results show why it is important to disentangle relative from absolute age. In this study, effects of absolute age go mostly in the opposite direction from those of relative age. For example, let us have a look at Tables B.3 and B.4 in the Appendix. The increase in absolute age by one year increases the chances of conducting risky health behaviors with respect to smoking and drinking as well as with respect to smocking marijuana and having had sex at least once in life, while it decrease the chances of having had sex without a condom during the last intercourse.

The results on the probability of having smoked tobacco or marijuana, and of having consumed alcohol before the age of 13, suggest a first possible mechanism behind the use of addictive substances. If students in the same class experiment with substances together, relatively younger students are mechanically more likely to initiate use earlier simply because they are part of the same cohort. This pattern of early initiation is important, as early tobacco use has been linked to a higher likelihood of persistent smoking (Bahrs & Schumann, 2020; Ellickson et al., 2001); this connection is reflected in the estimates for current smoking behavior in Table 3. Moreover, early consumption of substances, such as tobacco, alcohol, and marijuana has been associated with an increased likelihood of engaging in other risky behaviors later in life, such as binge drinking (Dai & Wang, 2023).

Do these effects last in time? We already have evidence that effects on health outcomes have the potential to last in time–at least on young adults, such as risky sexual behaviors (Johansen, 2021) and smoking Bahrs & Schumann (2020). However, to answer this question directly, with a comparable methodology, and within the context of risky health behaviors, we would need a dataset that: (a) has cross-country (and cutoff) variation; (b) is based on European countries; and (c) looks at either long-term consequences or, for example, long-term consumption of marijuana, tobacco, and alcohol. With all its limitations and only for illustrative purposes, *the Survey of Health, Ageing and Retirement in Europe (SHARE)* data could be a candidate dataset. However, the study on these data does not return reliable results.[25] Since we cannot consider this as the conclusive word on (very)long-term effects on risky behaviors, we encourage future studies to look more closely at this subject.

data preparation and analyses replication.

[25] We use SHARE data and investigate two outcome variables: (i) daily past smoking frequency during at least one year, and (ii) numbers of years smoking. There are no similar variables for alcohol–which focuses more on current drinking behaviors and thus is not informative (i.e., people who used to drink more frequently had higher probability of dying before the survey)–or marijuana consumption, for bullying or sexual behaviors. There are outcomes that proxy consequences of long-term or heavy consumption, such as liver and lungs cancer, but the number of occurrences does not grant enough statistical power. We run a reduced form, where we regress (i) and (ii) on, ERA, EAA, female, and fixed effects for wave, country, and season of birth. Since SHARE follows individuals over time, we consider only the latest observation per individual. The results are not statistically significant, and we are not inclined to consider them as very reliable, because of the dataset limitations and limited statistical power. While this paper does not include these results, the replication package includes the Stata do file for data preparation and analyses replication.

6 Potential Mechanisms

In this section, we investigate potential mechanisms in two ways. First, we conduct causal mediation analyses and, second, we investigate whether relative age affects perceived risks of consuming addictive substances and perceived peers' prevalence of consumption of such substances.

6.1 Causal Mediation Analyses

We explore five different channels through which relative age might affect risky behaviors: (a) student's academic self-concept in relation to her classmates; (b) perceived well-being; (c) satisfaction with own body image; (d) time spent outside with friends in the evening; and (e) perceived support from other students. We refer to these variables as *mediators* (denoted via M).

To assess the role of these five mechanisms, we conduct mediation analyses on five outcomes that pass the overidentification test in the main analysis and for which, thus, we are confident the instruments are valid. These outcomes are: (i) smoking, (ii) ever drunk, (iii) early marijuana, (iv) ever sex, (v) fight.[26]

The model specification is as follows:

$$\begin{aligned} RA_i &= \beta^0_{RA} + \beta^{ERA}_{RA} ERA_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{RA}} + \epsilon^{RA}_i, && (6)\\ M_i &= \beta^0_{M} + \beta^{RA}_{M} RA_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{M}} + \epsilon^{M}_i, && (7)\\ Y_i &= \beta^0_{Y} + \beta^{RA}_{Y} RA_i + \beta^{M}_{Y} M_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{Y}} + \epsilon^{Y}_i, && (8) \end{aligned}$$

where Equation 6 is the first stage of relative age, Equation 7 models the relationship between the mediator variable M, relative age, and additional covariates, while Equation 8 describes the outcome variable as a function of relative age, the mediator, and covariates.

In this section we follow Dippel et al. (2022)'s framework, which allows the usage of one instrument for both mediators and endogenous treatment. To closely follow the causal mediation analysis *à la Dippel et al.*, in these analyses expected relative age is not disaggregated and absolute age is treated as exogenous–we see in the Appendix that either change does not have any consequence on the estimates, returning identical results to those in Table 3.

The estimation procedure is as follows. The parameter β^{RA}_M is estimated via 2SLS, where the first stage is given by Equation 6 and the second stage is the following one:

$$M_i = \beta^0_{M} + \beta^{RA}_{M} \widehat{RA}_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{M}} + \epsilon^{M}_i, \qquad (9)$$

Here, $\widehat{RA}_i$ are the predicted values from the first stage in Equation 6 .

Parameters β^M_Y and β^{RA}_Y are estimated via 2SLS, where the mediator is instrumented with expected relative age and, importantly, RA_i is included as a regressor:

$$\begin{aligned} M_i &= \beta^0_{M} + \beta^{RA}_{M} RA_i + \beta^{ERA}_{M} ERA_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{M}} + \epsilon^{M}_i, && (10)\\ Y_i &= \beta^0_{Y} + \beta^{RA}_{Y} RA_i + \beta^{M}_{Y} \widehat{M}_i + \mathbf{X_i}\beta^{\mathbf{X}}_{\mathbf{Y}} + \epsilon^{Y}_i, && (11) \end{aligned}$$

Here, $\widehat{M}_i$ are the predicted values from the first stage in Equation 10.

[26]The overidentification test is passed also in the analyses on bully and cyberbullied, but we focus on fight because it is more policy relevant.

The *direct effect* is estimated with coefficient $\widehat{\beta_Y^{RA}}$. The *indirect effect* is the product of the estimated effect of RA_i on M_i, that is $\widehat{\beta_M^{RA}}$, and of the estimated effect of M_i on Y_i, that is $\widehat{\beta_Y^M}$. The *total effect* is the sum of the two: $\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}$.

In our analyses, we focus on the percentage of the indirect effect out of the total effect, that is:

$$\frac{\text{Indirect effect}}{\text{Total effect}} = \frac{\text{Indirect effect}}{\text{Direct effect} + \text{Indirect effect}} = \frac{\widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}}{\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}}. \tag{12}$$

These computations can be conducted with the Stata command `ivmediate` or step-by-step as in the supplementary material of (Dippel et al., 2022).[27]

We believe that the results from these mediation analyses could be causally interpreted. While we allow for unobserved confounders jointly affecting relative age and mediator ($\text{corr}(\epsilon^{RA}, \epsilon^M) \neq 0$) and unobserved confounders jointly affecting mediator and risky behaviors ($\text{corr}(\epsilon^M, \epsilon^Y) \neq 0$), the *key identifying assumption* is that there is no unobservable variable *orthogonal* to the mediator that affects both the outcome and relative age, while conditioning on family SES, parents being at home, season of birth, wave and country fixed-effects. More precisely, any relative age–risky behaviors confounder has to operate via the mediator and therefore $\text{corr}(\epsilon^{RA}, \epsilon^Y) = 0$.[28]

The following figure illustrates a causal structure that is compatible with this identifying assumption from Dippel et al. (2020).

*****Figure 1 about here*****

We cautiously believe that this key identifying assumption holds in our case. For example, consider a factor like a student's underlying health, which might influence both risky behaviors and relative age. A lower level of health could lead to delayed school entry, affecting relative age (the treatment). This, in turn, could impact risky behaviors (the outcome) later on. Furthermore, health conditions are unlikely to be orthogonal to any mediator we study. Consequently, underlying health can serve as an example of an unobserved confounder (perhaps the most important) that is compatible with the specific identifying structure that is necessary to make us of the setup of Dippel et al. (2020). Another standard confounder is family SES; however, our analyses control for that.[29]

Table 4 presents the results of the mediation analyses, including the total, direct, and indirect effects, as well as the proportion of the total effect accounted for by the indirect effect.[30]

[27]We also note that the total effect obtained by 2SLS estimation of the effect of RA_i on Y_i instrumented by ERA_i (see Equation 5, thus without any mediator, is algebraically identical to $\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}$ in the just identified case with a single instrument.

[28]See p.617 in Dippel et al. (2020). Formally $\epsilon^{RA} \perp \epsilon^Y | X$ needs to hold, while conditional dependence is allowed $\epsilon^{RA} \not\perp \epsilon^Y | \epsilon^M, X$.

[29]We should note that no statistical test that would determine the violation of the identifying assumption is currently available.

[30]Due to missing values on the mediators, the sample size of these analyses may vary from that seen in Table 4, but this has very little effect on the magnitude of the estimates. These additional analyses work as robustness checks too: the sample size is reduced due to mediators missing values, but the estimates are indistinguishable from the main ones. This suggests that unobservable characteristics of students who did not answer the mediators' questions are orthogonal to relative age.

*****Table 4 about here*****

The results are also visualized in the Appendix, Figure B.5, for all the combinations of outcomes and mediators.

Ratios close to one suggest that the indirect effect (i.e., the effect of relative age through the mediator, $\widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}$) accounts for the majority of the total effect (i.e., the net effect of relative age plus the indirect effect, $\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^M} \cdot \widehat{\beta_M^{RA}}$). Coefficients larger than one mean that the sign of the indirect effect is opposite to that of the direct effect. For example, more practically, the risky behavior response to increases in relative age would have been even stronger if solely the mediator was affected (Dippel et al., 2022).[31]

We can observe three main results. First, indirect effects on ever sex and fight are close to zero, suggesting that relative age affect these outcomes (almost) exclusively directly. Relatively older students are more likely to have had sex and to be involved in fights, simply because they are older than their classmates. It seems that simply because a student is older than the peers, this student is more likely to have had sex and to be involved in physical fights, net of age at school start, time in school, and age at outcome measuring.[32] Second, academic self-concept, well-being, body image, and students' support explain a large part of the total effect. Third, we expected that frequent evenings spent out with friends–which acts as a proxy for lower adults' supervision and thus greater autonomy with respect to health behaviors–would play an important role in channeling indirect effects, but we do not seem to observe this result.

These results could be causally interpreted, cautiously. We cannot think of any unobservable variable orthogonal to the mediators that affects both risky behaviors *and* relative age, while conditioning on the complete set of control variables. If such unobservable factors existed, our results would not carry causal interpretation. Also, note that, since we instrument the mediator with expected relative, this approach solves the possible reverse causality with Y affecting the mediator.

We can gain some interesting insights also from the effects of relative age on the mediators, reported in Table B.16 in the Appendix. The results suggest that the frequency at which students spend evenings out with friends is not affected by relative age; although we believe it is the first time this (lack of) relative age effect is studied, this result comes as a surprise. We expected that students who are perceived as being more mature than their peers would have enjoyed more autonomy. Moreover, relatively older students experience higher academic self-concept, well-being, and support from peers, as well established in the literature. Finally, relatively older students are more likely to have a positive view of their own body, as suggested in recent studies (Carpenter & Churchill, 2024; Fumarco et al., 2024).

We do not conduct one unique mediation analysis with all mediators due to interpretability reasons. While one may argue that this approach would allow the comparison of indirect effects, it would defy apriori the goal of gaining a causal interpretation: mediators could affect each other causing endogeneity. Moreover, it would not be possible to

[31] There are also a couple of negative ratios. In those cases, the mediator acts as a suppressor; it works in the opposite direction to the direct effect and it is larger in absolute terms. In this way, the denominator has opposite sign of the numerator.

[32] As mentioned earlier, the list of mediators we use is based on a survey of the literature and it is not meant to be complete. It is possible that there exist other mediators through which relative age affects these two risky behaviors. Thus, what we interpret as direct effect might actually be a combination of the actual direct effect and indirect effects through other mediators we are not testing.

conduct the causal mediation analysis *à la Dippel et al. (2022)*. However, one should note that Table 4 allows the comparison of indirect effects. In alternative, one could conduct the principal component analysis, as in Dippel et al. (2022). However, within the context of our study, this approach does not allow us to interpret meaningfully the results, as we have seen different mediators might imply very different indirect effects, in terms of both sign and magnitude–except when the outcomes are Ever Sex and Fight, which are impacted almost exclusively directly by relative age.[33]

6.2 Perceived Risk of Harming Themselves Self-harm and Perceived Peers' Consumption Prevalence

To investigate in greater detail the potential mechanisms underlying the effect of relative age on the consumption of addictive substances, we conduct additional analyses on data from the *European Monitoring Centre for Drugs and Drug Addiction (ESPAD)* survey. This survey is on 15-16 year-old students, from 28 European countries, and covers the period 1995-2015. The survey is conducted every four years since 1995; however, we could not use waves 1999 and 2003 because they do not include information on the month of birth. Similar to what we do in the main analyses with HBSC data, we exclude countries for which we do not have information on the cutoff date, the cutoff date falls in the middle of the month, or adopts multiple cutoffs across regions. Table B.17 in the Appendix reports the full list of countries included in this analyses, with their respective cutoff, and the number of observations per wave.[34] This table allows us to note the comparability of the two datasets: almost two-thirds of the countries in the ESPAD are also in the HBSC. These two surveys have additional points in common: the ESPAD's primary sampling unit is the class, many scientists are involved in the organization and study of both surveys, and ESPAD even coordinates some features with HBSC, such as the timing of the survey.[35]

For these analyses, we are unable to use the 2SLS and to conduct analyses at the class level due to two main reasons. First, with ESPAD data we cannot build a unique class identifier across wave, and we are not able to accurately create it either; thus, analyses here are at the cohort level. Second, due to the sensitivity of the topics being dealt with in the ESPAD survey, the sample size is quite smaller than the HBSC one. Therefore, in these secondary analyses, we focus on reduced-form effects, similar to the

[33] We note also that, within the context of our study, it would be misleading to conduct the principal component analysis with the purpose of estimating "the true" repartition of the total effect between direct and indirect effect. This approach implicitly assumes that we would be investigating all the possible mediators.

[34] Differently from the HBSC survey, the United Kingdom has one unique ESPAD survey, without specification on the state. Therefore, since cutoff dates vary across UK states, we cannot use data from there. Moreover, differently from the HBSC survey, Germany is divided in two parts: Bayern and the rest of Germany, without specification on the state. Therefore, since cutoff dates vary across states, we cannot use data from the rest of Germany, but we can use data from Bayern. Finally, observations from Russia, Albania and Moldova are excluded from the main analyses with HBSC data, because we do not unambiguously know their cutoff date, we do not have issues of statistical power, and we are concerned with precise unbiased estimates. Differently, in these reduced-form analyses, we are concerned with the statistical power, so we include these three countries despite the ambiguity about the actual cutoff date being used. Notice that data on Russian students come from Moscow only–which is where a large percentage of Russians live. Data from Belgium, Wallonia, do not include information from month of birth, and thus it is not included.

[35] As explicitly stated in the ESPAD website.

recent Oosterbeek et al. (2021), Cygan-Rehm & Westphal (2024), and Barabasch et al. (2024).

We adopt the model specification in Equation 13:

$$Y_i = \beta_Y^0 + \beta_Y^{ERA} ERA_i + \beta_Y^{EAA} EAA_i + \beta_Y^X \mathbf{X_i} + \epsilon_i^Y. \tag{13}$$

We regress the outcome on ERA and EAA,[36] a vector of covariates that included dummies for gender, living with both parents, and SES. Moreover, there is a vector of fixed effects for the calendar month of birth, survey wave, and country. Thus, Equation 13 is the reduced form version of Equation 5.

In these analyses, ERA is not disaggregated for two main reasons. First, the smaller sample size implies that there would be fewer observations per academic month dummy. Second, since we use the continuous version of ERA, we can normalize it and measure it in years (i.e., it ranges between zero and one), with one corresponding to being born in the month that starts with the cutoff. The convenient implication is that the interpretation of the reduced form estimate of ERA is similar to that of RAEs in the main analyses. With this reduced form, we investigate seven outcomes that are divided into two categories.

Perceived risk of harming yourself while consuming addictive substances. Outcomes in this category equal one if the student answered the question "How much do you think people risk harming themselves (physically or in other ways), if they ..." with "more than slight risk." These are the five outcomes being investigated: (1) smoke an occasional cigarette; (2) have 4-5 drinks per week; (3) occasionally smoke marijuana; (4) try ecstasy; and (5) try amphetamine.

Perceived peers' frequency of consumption of addictive substances. Outcomes in this category equal one if the student answered the question "How many of your friends would you estimate ..." with "more than a few." These are the seven outcomes being investigated: (1) smoke; (2) drink alcohol; (3) get drunk; (4) smoke marijuana; (5) uses tranquilizers; (6) uses ecstasy; and (7) uses inhaler.

Tables B.18 in the Appendix reports observations and means for the outcomes, while Table B.19 reports observations, means, and standard deviations, for expected relative and absolute age, and the other independent variables.

Results on perceived risk are shown in Table 5.

*****Table 5 about here*****

This table shows that relatively old students are between 1.2 and 2.5% more likely to think that smoking tobacco and marijuana, drinking alcohol, and trying ecstasy could be risky activities. Although these point estimates are small, they would be reflected by large absolute numbers, out of the universe of adolescents in a country. Differently, there is no evidence of a reduced form association between relative age and thinking that trying amphetamines could be a risky behavior.[37]

Results on perceived peers' frequency of consumption of addictive substances are shown in Table 6.

[36] Results obtained with observed absolute age in lieu of expected absolute age return virtually identical results.

[37] We conduct robustness checks without Russia, Albania and Moldova–for which we found ambiguous information on the adopted cutoff date–and obtain identical results, in terms of both magnitude and statistical significance; an exception is the result on smoke, which becomes statistically significant at 10% with a point estimate of -0.010.

*****Table 6 about here*****

This table shows that relatively old students are between 1.1 and 2.1% less likely to think that the consumption of various addictive substances is frequent among their peers. Differently, there is no relative age effect on perceived peers' consumption prevalence of tranquilizers and ecstasy, while the effect on inhaler is also negative and statistically significant, but only at 10%.

Overall, these findings indicate that relatively younger students tend to perceive addictive substances as less risky than their older peers do. They also believe that substance use is more common among their friends than older students think. If we assume these perceptions influence behavior, they represent additional mechanisms that may help explain why relatively younger students are more likely to use addictive substances. Specifically, younger students' higher consumption rates could partly stem from underestimating the risks of substance use. Additionally, they may overestimate how widespread these behaviors are among their peers, leading them to believe that substance use is more socially acceptable or common than it truly is.

7 Conclusion

This paper investigates the effects of relative age within classrooms on a comprehensive set of adolescents' risky health behaviors. In doing so, we contribute to two main strands of the economics literature: one examining the impact of relative age on risky behaviors and their determinants, and the other exploring the role of educational systems and peers in shaping individual risky behaviors.

Our main findings reveal distinct patterns across types of behaviors. Relatively younger students are more likely to smoke tobacco and marijuana and to consume alcohol, but less likely to have sex without a condom. In contrast, relatively older students are more likely to be involved in physical fights and to engage in bullying, while younger students are more often victims of bullying—but not of cyberbullying.

Causal mediation analyses suggest that academic self-concept, well-being, self-esteem, and peer support amplify the impact of relative age on substance use. In other words, these mediators appear to magnify the effect that relative age would have independently. By contrast, engagement in sexual activity and physical aggression appears to be driven more directly by relative age, with little mediation by these factors.

Taken together, these findings suggest a possible underlying mechanism. Risky behaviors related to sex and aggression—such as unprotected sex and fighting—may be more closely tied to physical development, with relatively older students more likely to engage in them. In contrast, for behaviors like substance use, where physical maturity may matter less, relatively younger students appear more vulnerable.

Reduced-form results from secondary analyses suggest additional mechanisms. Relatively younger students tend to underestimate the risks associated with addictive substances. They may also overestimate peer consumption or spend more time with peers who use these substances. These beliefs likely contribute to higher rates of substance use among younger students.

These findings suggest several important policy implications. First, targeted tutoring programs could help reduce performance gaps associated with relative age effects, which could reduce the mediator effect of students' academic self-concept. For instance, higher education students could be incentivized to tutor relatively younger peers, following the model proposed by Kraft & Falken (2021). Second, to address the heightened vulnerability of younger students to risky behaviors, schools could provide additional mental health and emotional support tailored to their specific needs. Third, informational campaigns aimed at younger students could help correct misperceptions about the prevalence and risks of substance use, potentially reducing the likelihood of early experimentation. Fourth, introducing "age allowances" to adjust for performance gaps could be a viable policy option. Although this approach has a long tradition, it is currently used only in specific areas of England (Peña, 2022). Its broader adoption would require careful discussion. For example, since relative age is a continuum, who should qualify for this allowance and would this allowance continuously change? Would allowances be granted regardless of socioeconomic status? These and other interventions may be relatively inexpensive while enhancing fairness in education and yielding potential public health benefits; however, they should be carefully designed and tested first.

Regardless of their cost, these interventions are also justifiable on the grounds of equality. The youngest students in a grade may be experiencing a form of "indirect discrimination," as defined in EU and some Anglo-Saxon legal systems: a seemingly neutral rule—such as grouping students by 12-month age bands—can systematically disadvantage individuals with certain demographic characteristics. In this case, depending on the type of risky behavior, the policy may disproportionately affect either the youngest or the oldest students in a class. [38]

To deepen our understanding of relative age effects—particularly their policy implications—future research should more systematically investigate the underlying mechanisms. Only a few studies have done so to date, including Page et al. (2019) and Page et al. (2017). For instance, to our knowledge, no study has examined the relationship between relative age and time-inconsistent preferences, which are known to contribute to poor consumption choices, including addiction and unhealthy diets (Gruber & Mullainathan, 2005; Read & Van Leeuwen, 1998).

[38]Greater details on indirect discrimination in the EU and Anglo-Saxon countries are provided, for example, in the Eurofund website, UK Statute Law Database website, and Australian Law Reform Commission website.

References

Agan, A., Doleac, J. L., & Harvey, A. (2023). Misdemeanor prosecution. *The Quarterly Journal of Economics*, *138*(3), 1453–1505.

Aizer, A., & Doyle Jr, J. J. (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics*, *130*(2), 759–803.

Angrist, J. D., & Frandsen, B. (2022). Machine labor. *Journal of Labor Economics*, *40*(S1), S97–S140.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, *106*(4), 979–1014.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics.* Princeton university press.

Argys, L. M., & Rees, D. I. (2008). Searching for peer group effects: A test of the contagion hypothesis. *Review of Economics and Statistics*, *90*(3), 442–458.

Arnold, G., & Depew, B. (2018). School starting age and long-run health in the united states. *Health Economics*, *27*(12), 1904-1920.

Averett, S. L., Argys, L. M., & Rees, D. I. (2011). Older siblings and adolescent risky behavior: does parenting play a role? *Journal of Population Economics*, *24*, 957–978.

Bahrs, M., & Schumann, M. (2020). Unlucky to be young? the long-term effects of school starting age on smoking behavior and health. *Journal of Population Economics*, *33*(2), 555–600.

Balestra, S., Sallin, A., & Wolter, S. C. (2021). High-ability influencers? the heterogeneous effects of gifted classmates. *Journal of Human Resources*, 0920–11170R1.

Ballatore, R. M., Paccagnella, M., & Tonello, M. (2020). Bullied because younger than my mates? the effect of age rank on victimisation at school. *Labour Economics*, *62*(C), S0927537119301083.

Barabasch, A., Cygan-Rehm, K., & Leibing, A. (2024). Timing of school entry and personality traits in adulthood. *IZA DP No. 17387*.

Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, *121*(4), 1437–1472.

Bhuller, M., Dahl, G. B., Løken, K. V., & Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, *128*(4), 1269–1324.

Black, S., Devereux, P., & Salvanes, K. G. (2011). Too young to leave the nest? the effects of school starting age. *The Review of Economics and Statistics*, *93*(2), 455-467.

Bound, J., & Jaeger, D. A. (2000). Do compulsory school attendance laws alone explain the association between quarter of birth and earnings? In *Research in labor economics.* Emerald Group Publishing Limited.

Buckles, K., & Hungerman, D. (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics*, *95*(3), 711-724.

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., ... Zuber, S. (2013). Data resource profile: the survey of health, ageing and retirement in europe (share). *International Journal of Epidemiology*, *42*(4), 992–1001. doi: 10.1093/ije/dyt088

Card, D., & Giuliano, L. (2013). Peer effects and multiple equilibria in the risky behavior of friends. *Review of Economics and Statistics*, *95*(4), 1130–1149.

Carpenter, C. S., & Churchill, B. F. (2024). Relative age, self-image, and weight-related health behaviors. *Working Paper*.

Cascio, E., & Schanzenbach, D. (2016). First in the class? age and the education production function. *Education Finance and Policy*, *11*(3), 225-250.

Clarke, D., Oreffice, S., & Quintana-Domeque, C. (2019). The demand for season of birth. *Journal of Applied Econometrics*, *34*(5), 707–723.

Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the health behaviour in school-aged children (hbsc) family affluence scale. *Social Science & Medicine*, *66*(6), 1429–1436.

Currie, J., & Schwandt, H. (2013). Within-mother analysis of seasonal patterns in health at birth. *Proceedings of the National Academy of Sciences*, *110*(30), 12265–12270.

Cygan-Rehm, K., & Westphal, M. (2024). School starting age and the gender pay gap over the life cycle. *LIfBi Working Paper No. 118*.

Dai, Z., & Wang, K. (2023). The association between early onset of alcohol, smokeless tobacco and marijuana use with adult binge drinking in united states. *Scientific Reports*, *13*(1), 187.

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review*, *25*(1), 43–62.

Dhuey, E. (2016). Age at school entry: How old is old enough? a child's age at school entry matters, and the implications of policy changes can have long-lasting effects. *IZA World of Labor*, *247*.

Dhuey, E., & Lipscomb, S. (2010). Disabled or young? relative age and special education diagnoses in schools. *Economics of Education Review*, *29*(5), 857–872.

Dickert-Conlin, S., & Elder, T. (2010). Suburban legend: School cutoff dates and the timing of births. *Economics of Education Review*, *29*(5), 826–841.

Dippel, C., Ferrara, A., & Heblich, S. (2020). Causal mediation analysis in instrumental-variables regressions. *The Stata Journal*, *20*(3), 613–626.

Dippel, C., Gold, R., Heblich, S., & Pinto, R. (2022). The effect of trade on workers and voters. *The Economic Journal*, *132*(641), 199–217.

Di Tella, R., & Schargrodsky, E. (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy*, *121*(1), 28–73.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, *108*(2), 201–240.

Duncan, T. E., Tildesley, E., Duncan, S. C., & Hops, H. (1995). The consistency of family and peer influences on the development of substance use in adolescence. *Addiction*, *90*(12), 1647–1660.

Dustmann, C., Sandner, M., & Schönberg, U. (2022). The effects of sun intensity during pregnancy and in the first 12 months of life on childhood obesity. *Journal of Human Resources*.

Ellickson, P. L., Tucker, J. S., & Klein, D. J. (2001). High-risk behaviors associated with early smoking: results from a 5-year follow-up. *Journal of Adolescent health*, *28*(6), 465–473.

Elsner, B., & Isphording, I. E. (2018). Rank, sex, drugs, and crime. *Journal of Human Resources*, *53*(2), 356–381.

Fletcher, J. M. (2012). Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach. *Journal of Population Economics*, *25*, 1265–1286.

Fredriksson, P., & Öckert, B. (2014). Life-cycle effects of age at school start. *The Economic Journal*, *124*(579), 977-1004.

Fumarco, L., & Baert, S. (2019). Relative age effect on european adolescents' social network. *Journal of Economic Behavior & Organization*, *168*, 318–337.

Fumarco, L., Baert, S., & Sarracino, F. (2020). Younger, dissatisfied, and unhealthy–relative age in adolescence. *Economics & Human Biology*, *37*, 100858.

Fumarco, L., Hartmann, S. A., & Principe, F. (2024). A neglected determinant of eating behaviors: Relative age. *IZA DP No. 16920*.

Fumarco, L., & Schultze, G. (2020). Does relative age make jack a dull student? evidence from students' schoolwork and playtime. *Education Economics*, *28*(6), 647–670.

Gans, J. S., & Leigh, A. (2009). Born on the first of july: An (un) natural experiment in birth timing. *Journal of Public Economics*, *93*(1-2), 246–263.

Gillen, M. M., Lefkowitz, E. S., & Shearer, C. L. (2006). Does body image play a role in risky sexual behavior and attitudes? *Journal of Youth and Adolescence*, *35*, 230–242.

Granner, M. L., Black, D. R., & Abood, D. A. (2002). Levels of cigarette and alcohol use related to eating-disorder attitudes. *American Journal of Health Behavior*, *26*(1), 43–55.

Gruber, J. H., & Mullainathan, S. (2005). Do cigarette taxes make smokers happier. *The BE Journal of Economic Analysis & Policy*, *5*(1).

Huang, C., Zhang, S., & Zhao, Q. (2020). The early bird catches the worm? school entry cutoff and the timing of births. *Journal of Development Economics*, *143*(C), S0304387818312082.

Johansen, E. R. (2021). Relative age for grade and adolescent risky health behavior. *Journal of Health Economics*, *76*, 102438.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review*, *96*(3), 863–876.

Kraft, M. A., & Falken, G. T. (2021). A blueprint for scaling tutoring and mentoring across public schools. *AERA Open*, *7*, 23328584211042858.

Landersø, R. K., Nielsen, H. S., & Simonsen, M. (2020). Effects of school starting age on the family. *Journal of Human Resources*, *55*(4), 1258–1286.

Lopez-Mayan, C., Montresor, G., & Nicodemo, C. (2024). School entry age policy and adolescent risk-taking. *IZA DP No. 17589*.

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. *American Economic Review*, *103*(5), 1797–1829.

Massey, E. K., Gebhardt, W. A., & Garnefski, N. (2008). Adolescent goal content and pursuit: A review of the literature from the past 16 years. *Developmental Review*, *28*(4), 421–460.

Mueller-Smith, M. (2015). The criminal and labor market impacts of incarceration. *Unpublished working paper*.

Mühlenweg, A. (2010). Young and innocent: international evidence on age effects within grades on victimization in elementary school. *Economics Letters*, *109*(3), 157–160.

Nicodemo, C., Nicoletti, C., & Vidiella-Martin, J. (2024). Starting school and adhd: When is it time to fly the nest? *IZA DP No. 17091*.

Oosterbeek, H., ter Meulen, S., & van der Klaauw, B. (2021). Long-term effects of school-starting-age rules. *Economics of Education Review*, *84*, 102144.

Page, L., Sarkar, D., & Silva-Goncalves, J. (2017). The older the bolder: Does relative age among peers influence children's preference for competition? *Journal of Economic Psychology*, *63*, 43–81.

Page, L., Sarkar, D., & Silva-Goncalves, J. (2019). Long-lasting effects of relative age at school. *Journal of Economic Behavior & Organization*, *168*, 166–195.

Pearl, J. (2009). *Causality*. Cambridge university press.

Peña, P. A. (2017). Creating winners and losers: Date of birth, relative age in school, and outcomes in childhood and adulthood. *Economics of Education Review*, *56*, 152–176.

Peña, P. A. (2020). Relative age and investment in human capital. *Economics of Education Review*, *78*, 102039.

Peña, P. A. (2022). End the birthday bias: Age allowances in high-stakes tests are a proven boost for fairness. *Education Next*, *22*(3), 22–29.

Peña, P. A., & Duckworth, A. L. (2018). The effects of relative and absolute age in the measurement of grit from 9th to 12th grade. *Economics of Education Review*, *66*, 183–190.

Read, D., & Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes*, *76*(2), 189–205.

Reynoso, A., & Rossi, M. A. (2019). Teenage risky behavior and parental supervision: the unintended consequences of multiple shifts school systems. *Economic Inquiry*, *57*(2), 774–791.

Routon, P. W., & Walker, J. K. (2022). (more) party time in school? relative age and alcohol consumption. *Applied Economics*, 1–17.

Santini, Z. I., Meilstrup, C., Hinrichsen, C., Nielsen, L., Koyanagi, A., Koushede, V., ... Madsen, K. R. (2020). Associations between multiple leisure activities, mental health and substance use among adolescents in denmark: A nationwide cross-sectional study. *Frontiers in behavioral neuroscience*, *14*, 593340.

Schwandt, H., & Wuppermann, A. (2016). The youngest get the pill: Adhd misdiagnosis in germany, its regional correlates and international comparison. *Labour Economics*, *43*, 72–86.

Sorbring, E., Skoog, T., & Bohlin, M. (2014). Adolescent girls' and boys' well-being in relation to online and offline sexual and romantic activity. *CyberPsychology*, *8*(1).

Sprietsma, M. (2010). Effect of relative age in the first grade of primary school on long-term scholastic results: international comparative evidence using pisa 2003. *Education Economics*, *18*(1), 1–32.

Springer, A., Parcel, G., Baumler, E., & Ross, M. (2006). Supportive social relationships and adolescent health risk behavior among secondary school students in el salvador. *Social Science & Medicine*, *62*(7), 1628–1640.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80.

Trudeau, J., Conway, K. S., & Kutinova Menclova, A. (2016). Soaking up the sun: The role of sunshine in the production of infant health. *American Journal of Health Economics*, *2*(1), 1–40.

Wernerfelt, N., Slusky, D. J., & Zeckhauser, R. (2017). Second trimester sunlight and asthma: Evidence from two independent studies. *American Journal of Health Economics*, *3*(2), 227–253.

Wild, L. G., Flisher, A. J., Bhana, A., & Lombard, C. (2004). Substance abuse, suicidality, and self-esteem in south african adolescents. *Journal of Drug Education*, *34*(1), 1–17.

Woertman, L., & Van den Brink, F. (2012). Body image and female sexual functioning and behavior: A review. *Journal of Sex Research*, *49*(2-3), 184–211.

Tables and figures

Table 1: Risky health behaviors.

Variable	N	Mean
Early smoking	308,506	0.241
Smoking currently	497,116	0.793
Early drinking	307,860	0.380
Ever drunk	480,690	0.255
Early marijuana	49,857	0.032
Ever marijuana	247,255	0.139
Ever sex	206,845	0.205
Unprotected sex	45,766	0.369
Fight	598,799	0.366
Bully	603,359	0.273
Bullied	603,444	0.282
Cyberbullied	242,366	0.143

Note: All these risky behaviors are dummy variables.

Table 2: Descriptive statistics for (expected) relative age, (expected) absolute age, control variables.

Variable	N	Mean	SD
Relative age (years)	597,327	-0.306	0.454
Absolute age (years, centered)	616,973	0.001	1.646
Expected relative age (months)	616,973	5.529	3.373
Expected absolute age (years)	616,973	13.521	1.635
Female	616,973	0.508	
Parents	596,387	0.760	
SES: Low	616,973	0.367	
SES: Medium	616,973	0.229	
SES: High	616,973	0.403	

Note: SES: Low is the reference dummy for the student's family socio-economic status. Female, Parents, and the SES variables are dichotomous variables; thus, the standard deviation (SD) is not reported. Analyses additionally include vectors for wave, country and season of birth.

Table 3: Relative age on all outcomes.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.024***	-0.069***	-0.050***
	(0.008)	(0.004)	(0.009)	(0.005)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014**	-0.022***	0.022**	0.055**
	(0.007)	(0.007)	(0.009)	(0.022)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.022***	0.002	0.041***
	(0.006)	(0.005)	(0.006)	(0.006)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. ***, **, * indicate significance at 1%, 5% and 10%, respectively. Full statistics, with estimates on control variables and ancillary tests, are reported in the Appendix in Tables B.3, B.4, B.5

Figure 1: Causal structure compatible with the identifying assumptions of Dippel et al. (2020), while exogenous variables are omitted for simplicity. Note that $\epsilon^R \perp \epsilon^Y | X$ holds because ϵ^M and Y are colliders on paths between ϵ^R and ϵ^Y. Conditioning on a collider ϵ^M opens up the $\epsilon^R \rightarrow \epsilon^M \leftarrow \epsilon^Y$ path and therefore $\epsilon^R \not\perp \epsilon^Y | \epsilon^M, X$, see e.g. Pearl (2009).

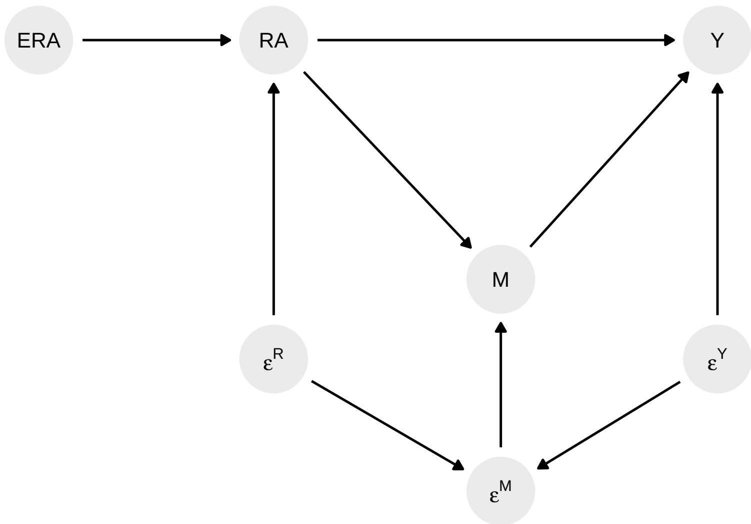


Table 4: Mediation analysis: Total ($\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^{M}} \cdot \widehat{\beta_M^{RA}}$), Direct ($\widehat{\beta_Y^{RA}}$), and Indirect Effects ($\widehat{\beta_Y^{M}} \cdot \widehat{\beta_M^{RA}}$) by Mediator for Different Outcomes. Indirect (%) = $\frac{\widehat{\beta_Y^{M}} \cdot \widehat{\beta_M^{RA}}}{\widehat{\beta_Y^{RA}} + \widehat{\beta_Y^{M}} \cdot \widehat{\beta_M^{RA}}}$ reports the proportion of the Total effect that is attributed to the Indirect effect.

Mediator variable	**Effect Type**	**Smoking**	**Ever Drunk**	**Early Marijuana**	**Ever Sex**	**Fight**
Academic self-concept	Total	-0.025	-0.049	-0.016	0.019	0.041
	Direct	-0.000	-0.022	-0.004	0.019	0.039
	Indirect	-0.024	-0.026	-0.012	0.001	0.002
	Indirect (%)	**98%**	**54%**	**75%**	**3%**	**5%**
Well-being	Total	-0.025	-0.051	-0.018	0.019	0.041
	Direct	0.007	-0.014	0.004	0.017	0.035
	Indirect	-0.032	-0.037	-0.021	0.002	0.006
	Indirect (%)	**129%**	**72%**	**122%**	**9%**	**14%**
Body image	Total	-0.024	-0.047	-0.018	0.023	0.043
	Direct	0.021	0.000	0.026	0.019	0.036
	Indirect	-0.045	-0.048	-0.044	0.004	0.006
	Indirect (%)	**187%**	**101%**	**241%**	**19%**	**15%**
Evening out	Total	-0.025	-0.050	-0.018	0.015	0.041
	Direct	-0.010	-0.033	-0.005	0.016	0.040
	Indirect	-0.015	-0.017	-0.013	-0.001	0.001
	Indirect (%)	**61%**	**34%**	**73%**	**-9%**	**2%**
Students' support	Total	-0.024	-0.049	-0.018	0.019	0.040
	Direct	0.018	-0.003	0.002	0.020	0.035
	Indirect	-0.043	-0.046	-0.019	-0.001	0.005
	Indirect (%)	**175%**	**94%**	**109%**	**-4%**	**13%**

Table 5: Relative age effects on perceived risk of harming themselves - reduced form.

	Occasional cigarette (1)	4-5 drinks a week (2)	Occasional marijuana (3)	Try ecstasy (4)	Try amphetamine (5)
ERA	0.016***	0.012***	0.025***	0.012**	0.003
	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)
N	121,572	120,480	113,768	106,354	105,922

Note: Outcome equals one if the answer to "How much do you think people risk harming themselves (physically or in other ways), if they . . . " is more than slight risk. ERA is not disaggregated and ranges between zero and one, with one corresponding to being born in the month that starts with the cutoff. Control variables are expected absolute age, female, both parents at home, medium SES, high SES, fixed effects for the month of birth within the calendar year, country, and wave. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table 6: Relative age effects on perceived peers' consumption prevalence - reduced form.

	Smoke (1)	Drink (2)	Drunk (3)	Marijuana (4)	Tranquilizers (5)	Ecstasy (6)	Inhaler (7)
ERA	-0.007	-0.011***	-0.021***	-0.018***	0.003	-0.002	-0.004*
	(0.006)	(0.004)	(0.006)	(0.004)	(0.002)	(0.002)	(0.002)
N	95,474	95,292	90,850	95,026	93,136	95,263	94,345

Note: Outcome equals one if the answer to "How many of your friends would you estimate . . . " is more than a few. ERA is not disaggregated and ranges between zero and one, with one corresponding to being born in the month that starts with the cutoff. Control variables are expected absolute age, female, both parents at home, medium SES, high SES, fixed effects for the month of birth within the calendar year, country, and wave. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

A Appendix

A.1 Expected relative age does not overlap with calendar month of birth

Figure A.1: Season of birth and expected relative age; example with Luxemburgish and Scottish students.

	Cal. year t												Cal. year t+1		
SOB	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar
ERA, Luxembourg	4	5	6	7	8	$\boxed{9}$	10	11	0	$\boxed{1}$	2	3	4	5	6
ERA, Scotland	10	11	0	1	2	3	4	5	6	7	8	9	10	11	0

Legend: Academic year x-1; Academic year x; Students' month of birth, Ex.1; Academic year x+1; Students' month of birth, Ex.2

Note: This figure illustrates two examples. Red cells illustrate Example 1. Here, there are two children born in September of the same calendar year t–and thus in the same season of birth, but in countries with different cutoff dates: September 1st in Luxembourg and March 1st in Scotland. Thus, they have different Expected Relative Age (ERA). Cells with the thick boarder illustrate Example 2. The thick-boarder cell under June shows that this student's ERA is 9, being born on the 9th month of academic year x-1 (i.e., light grey cells, before September), in Luxembourg. However, this student was retained and is placed in academic year x (i.e., darker grey cells, after September), where the oldest regular student is born four months later, in the thick-boarder cell under October. Thus, the retained student born in June is about four months older than the oldest regular student in the same class, who was born in October.

Example 1. This example illustrates why season of birth differs from ERA, for two students born in the same year and month, but in two countries with a different cutoff date. There are two children born in September of calendar year t; one student was born in Luxembourg, where the cutoff is September 1st, and the other one was born in Scotland, where the cutoff is March 1st. The Luxemburgish student is among the oldest students in their class since they were born in the month that starts with the cutoff date, while the Scottish student was born six months after the cutoff date month. These two students' months of birth are identified by red cells.

Because of differences in cutoff dates, the correlation between ERA and season of birth was about 0.549. This value does not cause problems of multicollinearity in the first stage. Table A.1 reports the variance inflation factors (VIFs) for both ERA and season of birth from the first stage: most of them are below 4.5, while the mean factor of the entire first stage is 3.62; these results are reassuring since the rule of thumb suggests that multicollinearity could be a problem when VIF>10.

Table A.1: Variance inflation factor of expected relative age and season of birth, from first stage.

Variables	VIF
ERA	
1	3.39
2	3.30
3	3.37
4	3.15
5	3.85
6	3.08
7	3.75
8	3.08
9	3.27
10	3.15
11	3.43
Season of birth	
1	4.03
2	3.95
3	4.18
4	3.51
5	4.62
6	3.81
7	4.46
8	3.35
9	4.08
10	3.69
11	4.18
Mean VIF	3.62

Note: VIFs' in this table, including the mean VIF, refer to the first stage.

Example 2. Rather than comparing ERA to season of birth, this example illustrates why RA differs from ERA, for two students born in the same year and country, but different month, and where one student was retained. Here, we consider a retained Luxemburgish student born in June, in the thick-boarder cell. They should be relatively young–being born in the ninth month of the academic year, but they are relatively old because of their retention. Assume that, in their current class, the oldest regular classmate (i.e., neither retained nor redshirted) was born in October, in the other thick-boarder cell. The retained student's ERA is nine because they were born nine months after the Luxemburgish cutoff date month; however, their (observed) RA is about 0.33, because they are about four months older than the actual oldest regular student in their class (i.e., the regular student born in October, in the thick-boarder cell).

Additionally, this second example intuitively illustrates why studies on age at school start capture the effect of both proper age at school start and relative age. Let us go on with Example 2 and assume that the two students are regular ones, so their ERA equals their RA. The student born in October has about the same age at the student born in June at the moment of the survey; however, these two students started school on two different academic years (the student born in June one year earlier than the student born on October) and are on the two opposite sides of their class age range.

A.2 Balance Tests on Expected Relative Age

There is one underlying assumption for using ERA as instrument, which is in common with most of the other literature on relative age and age at school start: birthdate has to be orthogonal to demographic variables (Dickert-Conlin & Elder, 2010).

Some studies find that families with different SES might target different dates of delivery. Evidence of this phenomenon, its magnitude, and its generalizability are a matter of debate. It has been found in Anglo-Saxon countries (e.g., in the US (Clarke et al., 2019; Buckles & Hungerman, 2013) and in Australia (Gans & Leigh, 2009)) and in the vicinity of the cutoff date (e.g., as discussed in Dhuey & Lipscomb (2010), through cesarean section, which in many countries is restricted to medical motivations), and with limited economic significance elsewhere (e.g., in China (Huang et al., 2020)). In general, this relationship depends on local characteristics, such as norms and tax incentives, as explained in Dickert-Conlin & Elder (2010).

We address the concern on the possible non-orthogonality of birthdate and demographic characteristics in three distinct ways. First, we conduct robustness checks where we replicate the main analyses, but we exclude family SES from the set of control variables; the main results are unchanged (see Table B.7). Second, we conduct robustness checks where we replicate the main analyses, but we exclude the control variable for having both parents at home; the main results are the same (see Table B.8). Third, we conduct joint orthogonality tests on the instrumental variable for relative age. In other words, we test the orthogonality of ERA with respect to observable demographic characteristics with a series of conditional balance tests, where we run OLS regressions of each demographic characteristic on the set of dummies for ERA and on fixed effects for country, wave, and season of birth. The results are reported in Table A.2 and they suggest that in Europe, on average, families with a different SES do not target different dates of birth. In other words, these results are reassuring as they suggest that ERA is randomly distributed with respect to observable characteristics. Most importantly, ERA is balanced with respect to parents' SES; this result rules out the possibility that parents tend to target certain birth dates depending on their SES–when we control for between-country differences with country fixed effects. These results on the unbiased nature of ERA, and thus on birth date exogeneity, suggest two things: (i) EAA is exogenous too, and (ii) absolute age, although theoretically endogenous, might be treated as exogenous[39]. This insight will be useful for conducting the mediation analyses à la Dippel et al. (2022). Analyses where we do not instrument absolute age return results that are equivalent to the main ones.

As a final note, this result does not call into question results from the above literature on birthdate targeting: these balance tests are conducted with country fixed-effects and thus are accounting for possible between-countries differences.

[39]That is, the possible bias in relative age effects estimates obtained when not instrumenting absolute age as well is so small that it does not make any economic difference

Table A.2: Conditional correlation between expected relative age and main control variables.

Variables	Female	Parents	SES: Low	SES: Medium	SES: High
	(1)	(2)	(3)	(4)	(5)
ERA 1	0.000	-0.003	-0.001	0.001	0.000
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
ERA 2	0.005	0.001	-0.000	-0.001	0.001
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
ERA 3	0.002	-0.002	-0.005	0.002	0.004
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
ERA 4	0.004	-0.001	-0.001	-0.002	0.003
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
ERA 5	0.001	-0.001	0.002	-0.000	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ERA 6	0.004	0.005	0.000	-0.003	0.002
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
ERA 7	0.010**	0.002	0.003	-0.001	-0.002
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
ERA 8	0.006	0.001	-0.000	0.002	-0.001
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
ERA 9	0.004	0.001	-0.002	0.003	-0.002
	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
ERA 10	0.003	-0.005	0.003	0.001	-0.004
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)
ERA 11	0.003	-0.004	0.004	-0.000	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
N	616,973	596,387	616,973	616,973	616,973

Note: ERAs are dummies for academic month of birth, that is, expected relative age. SES is socio-economic status. Clustered standard errors at the class level in parentheses.

B Appendix - Additional Tables

B.1 Observations by country and wave, and country cutoff date – HBSC data

Table B.1: Cutoff dates, and quantity of observations per country per wave, for the HBSC dataset.

		Wave					
Country	Cutoff date	2001/2	2005/6	2009/10	2014/15	2017/18	All waves
Austria	Sep 1st	4,150	4,757	4,679	3,313	3,794	20,693
Belgium, Flanders	Jan 1st	1,345	3,113	3,029	3,230	3,013	13,730
Belgium, Wallonia	Jan 1st	3,026	3,589	3,080	4,845	4,146	18,686
Bulgaria	Jan 1st	0	4,826	0	4,639	3,199	12,664
Croatia	Apr 1st	4,270	4,680	6,058	5,507	4,687	25,202
Czech Republic	Sep 1st	5,006	0	4,324	5,041	11,265	25,636
Denmark	Jan 1st	4,468	5,319	3,921	3,784	3,112	20,604
England	Sep 1st	3,822	4,697	3,437	5,261	3,084	20,301
Estonia	Oct 1st	3,279	4,202	4,131	4,001	4,592	20,205
Finland	Jan 1st	5,143	5,143	6,496	5,810	0	22,592
France	Jan 1st	7,393	5,736	5,457	5,168	8,599	32,353
Greece	Jan 1st	0	0	4,808	4,078	3,807	12,693
Greenland	Jan 1st	0	0	198	141	556	895
Hungary	Jul 1st	3,985	3,450	4,569	3,737	3,456	19,197
Iceland	Jan 1st		8,480	8,747	9,160	3,643	30,030
Ireland	Jan 1st	1,951	3,730	1,859	3,366	3,120	14,026
Italy	Jan 1st	4,313	3,867	4,734	3,906	4,025	20,845
Latvia	Jan 1st	3,225	4,096	4,053	4,924	3,946	20,244
Lithuania	Jan 1st	5,586	5,575	5,221	0	1,507	17,889
Luxembourg	Sep 1st	0	2,889	2,968	2,192	2,315	10,364
Malta	Jan 1st	1,853	0	0	2,227	1,936	6,016
Netherlands	Oct 1st	3,778	3,796	4,076	3,862	4,206	19,718
North Macedonia	Jan 1st	3,593	4,749	3,434	4,096	4,072	19,944
Norway	Jan 1st	4,943		4,050	3,144	2,891	15,028
Poland	Jul 1st	6,245	5,475	4,190	4,068	4,953	24,931
Scotland	Mar 1st	4,381	6,130	6,668	5,672	4,799	27,650
Slovakia	Sep 1st	0	0	4,468	4,997	0	9,465
Slovenia	Jan 1st	3,894	5,070	5,322	4,795	5,574	24,655
Spain	Jan 1st	5,418	7,738	3,861	3,442	2,938	23,397
Sweden	Jan 1st	3,778	4,332	6,627	7,471	4,076	26,284
Ukraine	Jan 1st	3,943	4,859	5,345	3,095	5,263	22,505
Wales	Sep 1st	3,771	4,384	5,326	5,050	0	18,531
All countries		106,559	124,682	135,136	134,022	116,574	616,973

Note: Flanders and Wallonia as well as Denmark mainland and Greenland hold separate surveys within Belgium and Denmark, respectively.

B.2 First stage results

Table B.2: First stage.

Variables	RA	AA
	(1)	(2)
ERA 1	0.001	-0.021***
	(0.004)	(0.003)
ERA 2	-0.040***	-0.062***
	(0.003)	(0.002)
ERA 3	-0.078***	0.037***
	(0.004)	(0.002)
ERA 4	-0.154***	-0.004**
	(0.003)	(0.002)
ERA 5	-0.205***	-0.040***
	(0.004)	(0.002)
ERA 6	-0.198***	0.043***
	(0.004)	(0.002)
ERA 7	-0.243***	0.011***
	(0.004)	(0.002)
ERA 8	-0.322***	-0.043***
	(0.005)	(0.002)
ERA 9	-0.346***	0.020***
	(0.005)	(0.003)
ERA 10	-0.336***	0.010***
	(0.005)	(0.003)
ERA 11	-0.341***	-0.004
	(0.006)	(0.003)
EAA	0.015***	0.990***
	(0.001)	(0.001)
N	577,691	596,387

Note: Clustered standard errors at the level of the class in parentheses. ERA stands for expected relative age. EAA stands for expected absolute age. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

B.3 Complete main results

Table B.3: RAE on smoking and drinking.

	Smoking		Drinking	
Variables	Early	Currently	Early	Ever drunk
	(1)	(2)	(3)	(4)
RA	-0.044***	-0.024***	-0.069***	-0.050***
	(0.008)	(0.004)	(0.009)	(0.005)
AA	0.053***	0.042***	0.042***	0.102***
	(0.001)	(0.000)	(0.001)	(0.001)
Female	-0.046***	-0.005***	-0.075***	-0.046***
	(0.002)	(0.001)	(0.002)	(0.002)
Parents	-0.080***	-0.047***	-0.060***	-0.076***
	(0.002)	(0.001)	(0.002)	(0.002)
SES: Medium	-0.006***	-0.009***	0.022***	0.006***
	(0.002)	(0.001)	(0.003)	(0.002)
SES: High	-0.005**	-0.009***	0.038***	0.021***
	(0.002)	(0.001)	(0.002)	(0.002)
N	285,742	464,318	285,190	448,722
Ancillary tests				
Underid. test: Lagrange	447.8	377.4	444.1	375.4
multiplier st. [p-value]	[0.001]	[0.001]	[0.001]	[0.001]
Weak id.: F-statistics	432.1	740.4	441.3	723.7
Overid. test: Hansen	17	7.780	17.51	5.590
J st. [p-value]	[0.074]	[0.650]	[0.064]	[0.848]

Note: Second stage estimates from the 2SLS. RA is relative age, AA is absolute age and it is centered around the mean to accommodate a more meaningful interpretation. SES is socio-economic status. Control variables include additionally include vectors for wave, country, and season-of-birth fixed effects. Clustered standard errors at the level of class in parentheses. Underid. stands for underidentification; Lagrange multiplier st. stands for Lagrange multiplier statistic; weak id. stands for weak identification; Overid. stands for overidentification; Hansen J st. stands for Hansen J statistics. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.4: RAE on marijuana consumption and sex.

	Marjuana		Sex	
Variables	Early	Ever	Ever	Unprotected
	(1)	(2)	(4)	(5)
RA	-0.014**	-0.022***	0.022**	0.055**
	(0.007)	(0.007)	(0.009)	(0.022)
AA	0.008***	0.060***	0.082***	-0.094***
	(0.001)	(0.001)	(0.002)	(0.009)
Female	-0.015***	-0.043***	-0.049***	0.060***
	(0.002)	(0.002)	(0.002)	(0.005)
Parents	-0.021***	-0.072***	-0.095***	0.027***
	(0.002)	(0.002)	(0.003)	(0.005)
SES: Medium	-0.001	0.008***	0.009***	-0.030***
	(0.002)	(0.002)	(0.003)	(0.006)
SES: High	0.001	0.017***	0.025***	-0.061***
	(0.002)	(0.002)	(0.002)	(0.006)
N	45,247	230,232	194,101	42,477
Ancillary tests				
Underid. test: Lagrange	194.7	344.6	387.3	327.9
multiplier st. [p-value]	[0.001]	[0.001]	[0.001]	[0.001]
Weak id.: F-statistics	160.9	308.2	370.5	181.9
Overid. test: Hansen	11.74	22.03	9.018	22.05
J st. [p-value]	[0.303]	[0.015]	[0.530]	[0.015]

Note: Second stage estimates from the 2SLS. RA is relative age, AA is absolute age and it is centered around the mean to accommodate a more meaningful interpretation. SES is socio-economic status. Control variables include additionally include vectors for wave, country, and season-of-birth fixed effects. Clustered standard errors at the level of the class in parentheses. Underid. stands for underidentification; Lagrange multiplier st. stands for Lagrange multiplier statistic; weak id. stands for weak identification; Overid. abbreviates overidentification; Hansen J st. stands for Hansen J statistics. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.5: RAE on aggressive behaviors.

Variables	Bullied (1)	Bully (2)	Cyberbullied (3)	Fight (4)
RA	-0.012**	0.022***	0.002	0.041***
	(0.006)	(0.005)	(0.006)	(0.006)
AA	-0.023***	0.009***	0.002***	-0.024***
	(0.000)	(0.000)	(0.001)	(0.001)
Female	-0.030***	-0.121***	0.029***	-0.300***
	(0.001)	(0.001)	(0.002)	(0.002)
Parents	-0.041***	-0.040***	-0.037***	-0.060***
	(0.001)	(0.001)	(0.002)	(0.002)
SES: Medium	-0.015***	0.004***	-0.002	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
SES: High	-0.023***	0.019***	-0.001	0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
N	566,168	566,196	224,131	561,632
Ancillary tests				
Underid. test: Lagrange	372.9	372.1	250.5	376
multiplier st. [p-value]	[0.001]	[0.001]	[0.001]	[0.001]
Weak id.: F-statistics	809.2	828	354	796.6
Overid. test: Hansen	20.47	13.54	5.008	4.373
J st. [p-value]	[0.025]	[0.195]	[0.891]	[0.929]

Note: Second stage estimates from the 2SLS. RA is relative age, AA is absolute age and it is centered around the mean to accommodate a more meaningful interpretation. SES is socio-economic status. Control variables include additionally include vectors for wave, country, and season-of-birth fixed effects. Clustered standard errors at the level of the class in parentheses. Underid. stands for underidentification; Lagrange multiplier st. stands for Lagrange multiplier statistic; weak id. stands for weak identification; Overid. stands for overidentification; Hansen J st. stands for Hansen J statistics. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

B.4 Robustness checks

B.4.1 Various model specifications

Model specifications used in this Appendix subsection B.4 are the same as in Equation 5, except for what is mentioned in the caption.

Table B.6: Relative age on all outcomes, continuous expected relative age as instrument.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.042***	-0.025***	-0.070***	-0.049***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.017**	-0.021***	0.022**	0.046**
	(0.007)	(0.007)	(0.009)	(0.022)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.014**	0.020***	0.003	0.041***
	(0.006)	(0.005)	(0.007)	(0.006)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.7: Relative age on all outcomes, without controlling for family socio-economic status.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.024***	-0.069***	-0.049***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014*	-0.022***	0.022**	0.055***
	(0.007)	(0.008)	(0.009)	(0.021)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.022***	0.002	0.041***
	(0.005)	(0.005)	(0.007)	(0.006)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.8: Relative age on all outcomes, without controlling for having both parents at home.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.024***	-0.069***	-0.049***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014*	-0.022***	0.022**	0.055***
	(0.007)	(0.008)	(0.009)	(0.021)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.022***	0.002	0.041***
	(0.005)	(0.005)	(0.007)	(0.006)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.9: Relative age on all outcomes, with class-level controls: class size, share of female students, and share of students with high SES.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.024***	-0.069***	-0.050***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014*	-0.022***	0.021**	0.056***
	(0.007)	(0.008)	(0.009)	(0.021)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.021***	0.001	0.040***
	(0.005)	(0.005)	(0.007)	(0.006)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the country level in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.10: Relative age on all outcomes, with controls for number of siblings at home and at least one of them being older.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.022***	-0.068***	-0.055***
	(0.011)	(0.006)	(0.013)	(0.008)
N	149,173	225,777	150,516	218,999
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-	-0.022*	0.021	0.054*
	-	(0.012)	(0.013)	(0.028)
N	-	97,272	79,320	19,197
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012	0.027***	-	0.052***
	(0.009)	(0.009)	-	(0.009)
N	224,352	224,405	-	221,263

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. The analyses with number of siblings at home and at least one of them being older (we added a semi-interaction term, i.e., $siblingsNUMBER + siblingsNUMBER * oneOLDER$) could not be conducted for early marijuana and cyberbulling because they were asked on different waves (i.e. the question on number of siblings is not asked in 2018, while the question on having at least one older sibling is not asked either in 2014 or 2018; the question on early marijuana was asked only in 2014 and the question on cyberbulling was asked only on 2014 and 2018). Clustered standard errors at the country level in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.11: Relative age on all outcomes, with standard errors clustered at country level.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.024***	-0.069***	-0.049***
	(0.011)	(0.004)	(0.017)	(0.008)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014	-0.022*	0.022	0.055***
	(0.011)	(0.012)	(0.014)	(0.036)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012	0.022***	0.002	0.041***
	(0.007)	(0.008)	(0.007)	(0.007)
N	566,168	566,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the country level in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.12: Relative age on all outcomes, with school instead of country fixed-effects.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.041***	-0.019***	-0.058***	-0.057***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,571	464,314	285,010	448,711
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.012	-0.016*	0.030***	0.083***
	(0.008)	(0.008)	(0.010)	(0.028)
N	45,218	230,134	194,000	40,995
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.011*	0.020***	0.001	0.042***
	(0.006)	(0.005)	(0.007)	(0.006)
N	566,168	566,159	224,123	561,620

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.13: Relative age on all outcomes, without extreme values for relative age (i.e., $RA <= -2$ or $RA > 1$).

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.046***	-0.025***	-0.070***	-0.050***
	(0.008)	(0.004)	(0.009)	(0.006)
N	282,106	457,103	281,558	441,925
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.014*	-0.023***	0.021**	0.056***
	(0.008)	(0.008)	(0.009)	(0.025)
N	44,716	226,563	191,315	41,780
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.011**	0.022***	0.002	0.042***
	(0.006)	(0.005)	(0.007)	(0.006)
N	557,053	557,135	220,275	552,927

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

Table B.14: Relative age on all outcomes, without instrumenting absolute age.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.044***	-0.025***	-0.069***	-0.050***
	(0.008)	(0.004)	(0.009)	(0.006)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.015*	-0.023***	0.021**	0.076***
	(0.008)	(0.008)	(0.010)	(0.028)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.022***	0.001	0.041***
	(0.006)	(0.005)	(0.007)	(0.006)
N	566,168	556,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

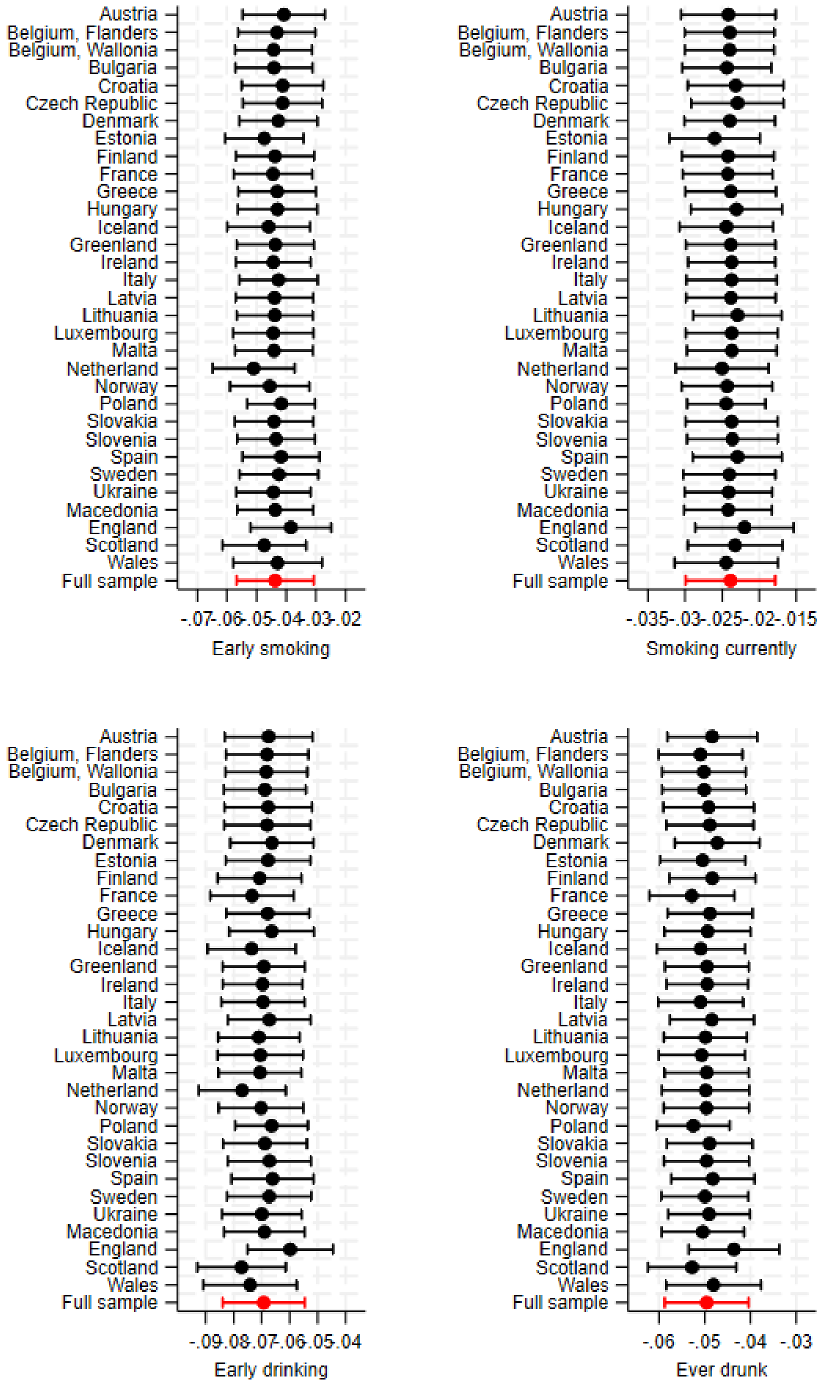
Table B.15: Relative age on all outcomes, with absolute age instrumented with the median of classmates absolute age.

	(1)	(2)	(3)	(4)
	Early smoking	Smoking currently	Early drinking	Ever drunk
Relative age	-0.043***	-0.024***	-0.069***	-0.049***
	(0.008)	(0.004)	(0.009)	(0.005)
N	285,742	464,318	285,190	448,722
	Early marijuana	Ever marijuana	Ever sex	Unprotected sex
Relative age	-0.015*	-0.023***	0.021**	0.076***
	(0.008)	(0.008)	(0.010)	(0.028)
N	45,247	230,232	194,101	42,477
	Bullied	Bully	Cyberbullied	Fight
Relative age	-0.012**	0.022***	0.001	0.041***
	(0.006)	(0.005)	(0.007)	(0.006)
N	566,168	556,196	224,131	561,632

Note: Second stage estimates from the 2SLS. Early smoking equals one if the student has already smoked once by 13 years of age. Smoking equals one if the student smokes at least once a week. Early drinking equals one if the student has drunk at least once by 13 years of age. Ever drunk equals one if the student has been drunk at least once in life. Early Marijuana equals one if the student has smoked marijuana at least once by 13 years of age. Ever smoked marijuana equals one if the student has smoked marijuana at least once in life. Ever sex equals one if the student has had sex at least once in life. Unprotected sex equals one if the student has had sex without a condom during the last intercourse. Fight equals one if the student was involved in at least one fight in the past year. Bully equals one if the student has bullied someone at least once during the past two months. Been bullied equals one if the student has been bullied at least once during the past two months. Cyberbullied equals one if the student has been cyberbullied at least once in life. Clustered standard errors at the level of the class in parentheses. ***, **, * indicate significance at 1%, 5% and 10%, respectively.

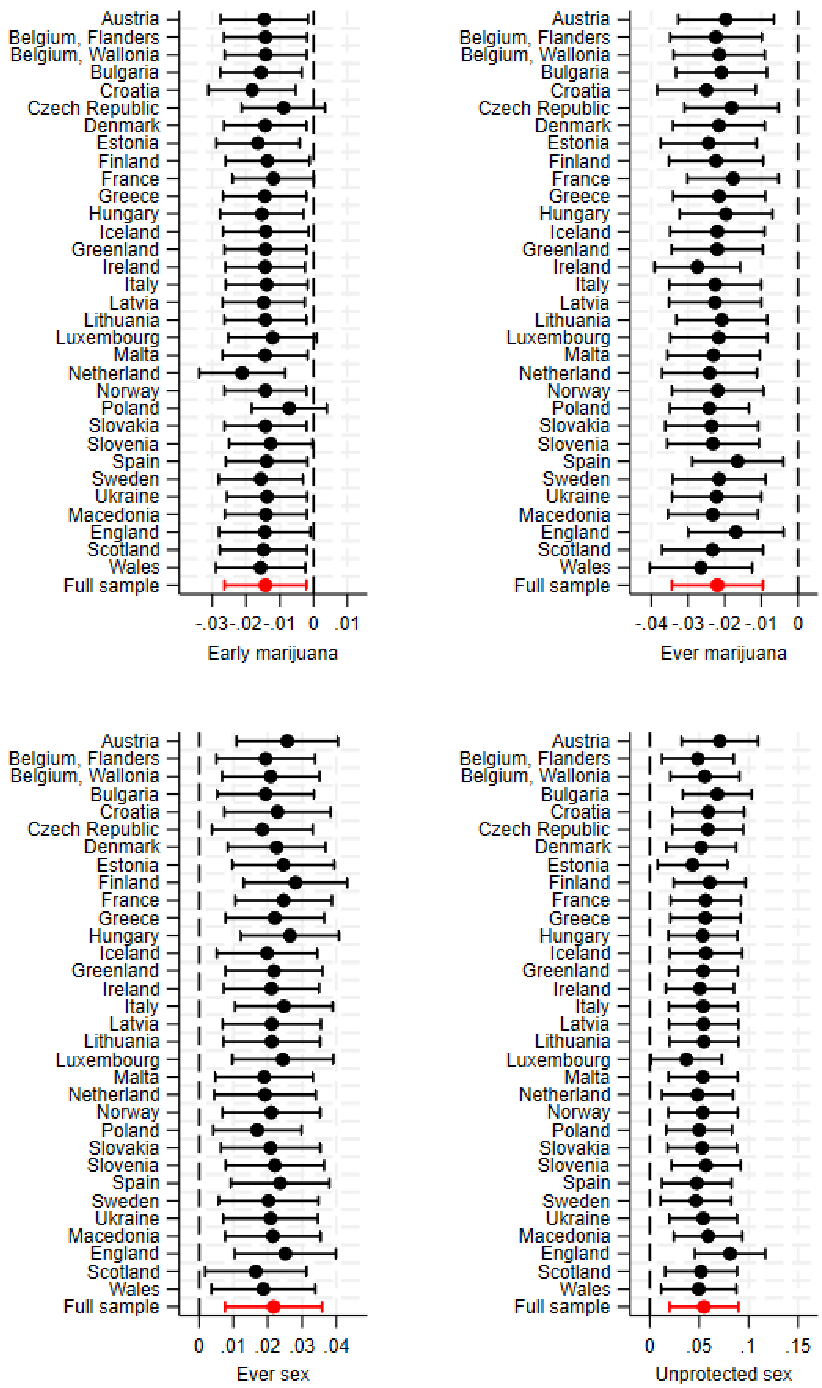
B.4.2 Leave-one-out

Figure B.1: Leave-one-out, Smoking and drinking



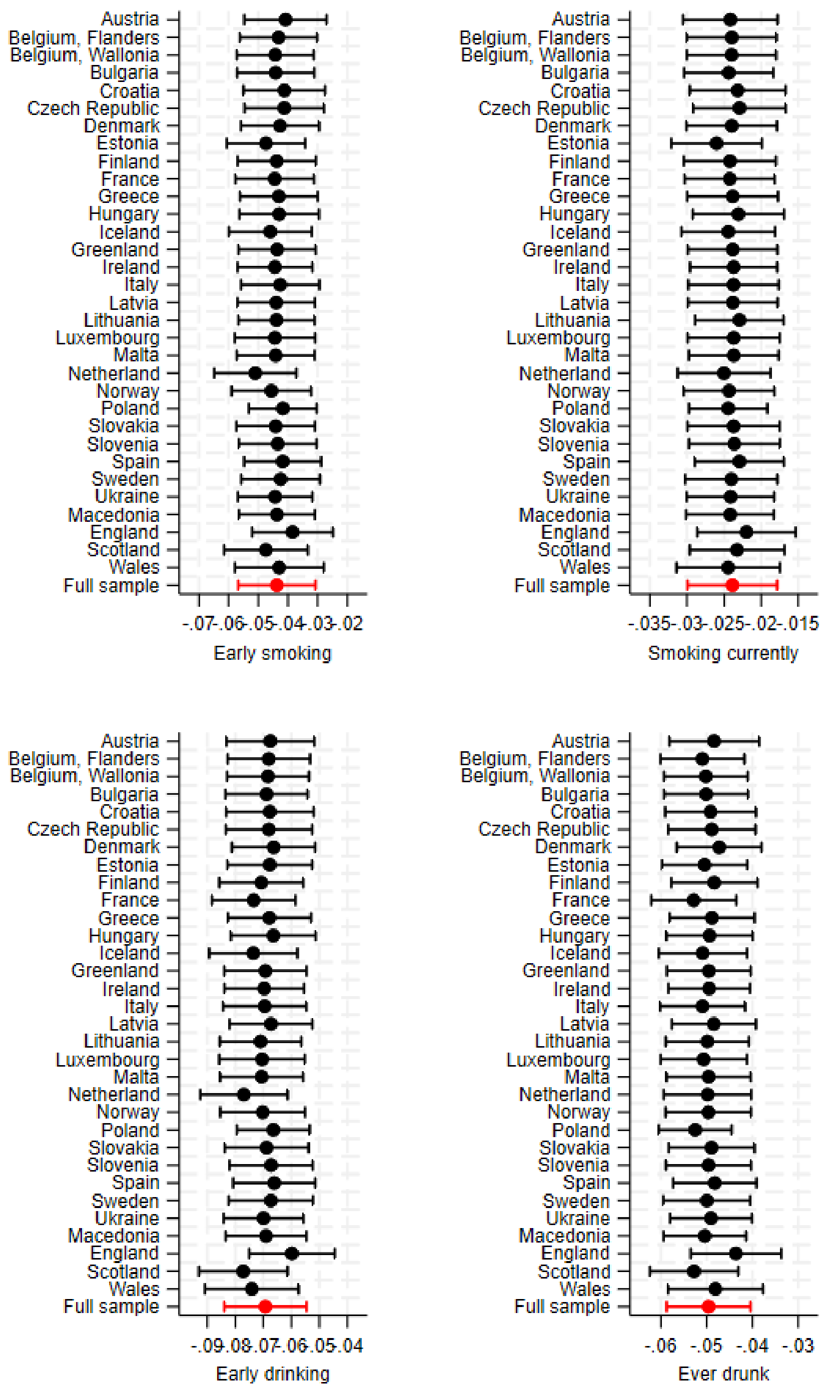
Note: The y-axis reports name of the country being left out of the analysis. The x-axis reports estimated relative age effects. The model specification is the same to Equation 5. 90% confidence intervals are reported.

Figure B.2: Leave-one-out, Marijuana and sex



Note: The y-axis reports name of the country being left out of the analysis. The x-axis reports estimated relative age effects. The model specification is the same to Equation 5. 90% confidence intervals are reported.

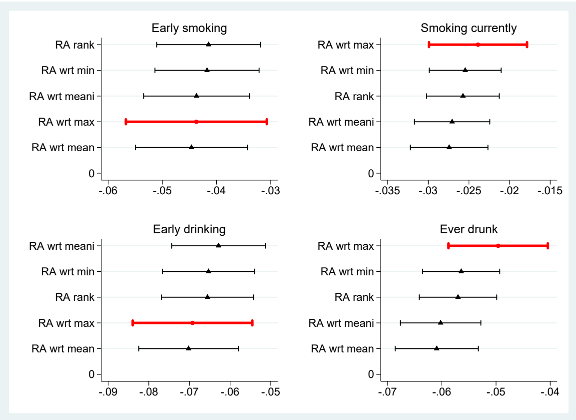
Figure B.3: Leave-one-out, Fighting



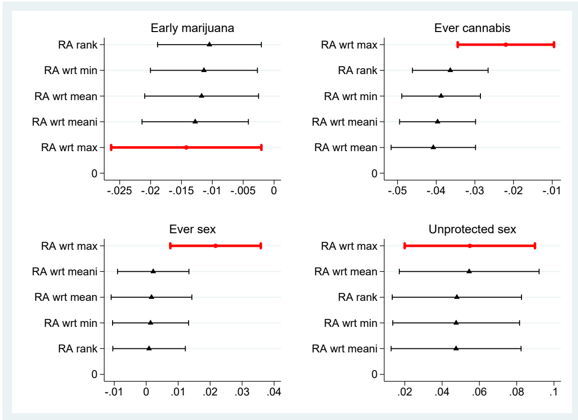
Note: The y-axis reports name of the country being left out of the analysis. The x-axis reports estimated relative age effects. The model specification is the same to Equation 5. 90% confidence intervals are reported.

B.4.3 Various relative age specifications

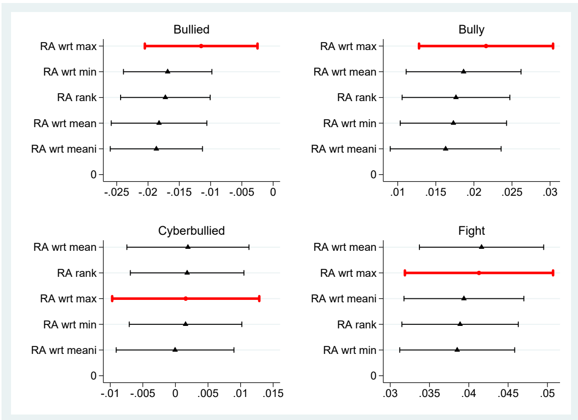
Figure B.4: Various relative age specifications, with the main model specification in Equation 5.



(a) Smoking and drinking



(b) Marijuana and sex



(c) Fighting

Note: The x-axis reports different relative age specifications, whereas the y-axis reports estimated relative age effects. The model specification is similar to Equation 5, except that the outcome varies: (i) "RA wrt max" is the measure of relative age used in all the analyses and that takes the difference between student i's age and that of the oldest regular student in the class; (ii) "RA wrt min" takes the difference with respect to the youngest regular student in the class; (iii) "RA wrt mean" takes the difference with respect to the mean age in student i's class; (iv) "RA wrt meani" the same as (iii), but the mean excludes student i's age; (v) "RA rank" is student i's percentile age rank within her class. 90% confidence intervals are reported.

B.5 Complete mediation analyses

Figure B.5: Results from mediation analysis: Total, Direct, and Indirect Effects by Mediator for Different Outcomes, see Table 4.

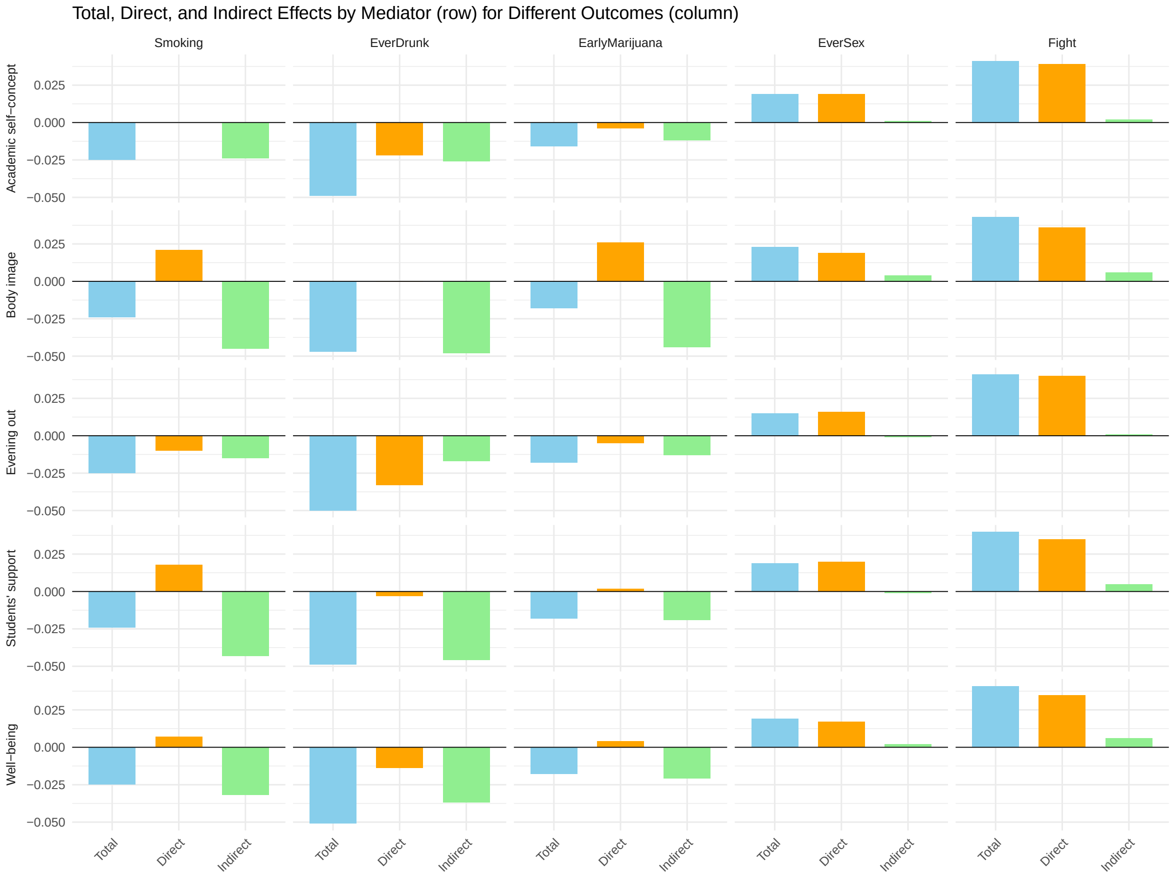


Table B.16: Effect on the mediator

Outcome Variable	Mediator	$RA \to M$, $\widehat{\beta_M^{RA}}$
Currently Smoking	Academic Self-Concept	0.213***
	Well-being	0.114***
	Body Image	0.040***
	Evening Out	0.023*
	Students' Support	0.125***
Ever Drunk	Academic Self-Concept	0.210***
	Well-being	0.117***
	Body Image	0.041***
	Evening Out	0.018
	Students' Support	0.130***
Early Marijuana	Academic Self-Concept	0.188***
	Well-being	0.123***
	Body Image	0.027
	Evening Out	0.030
	Students' Support	0.126***
Ever Sex	Academic Self-Concept	0.175***
	Well-being	0.115***
	Body Image	0.037***
	Evening Out	0.052*
	Students' Support	0.109***
Fight	Academic Self-Concept	0.207***
	Well-being	0.121***
	Body Image	0.045***
	Evening Out	0.021
	Students' Support	0.118***

Note: These results come from the 2SLS regression of the mediator on the treatment. Clustered standard errors at the class level in parentheses. ***, **, * indicate significance at 1%, 5%, and 10%, respectively.

B.6 Observations by country and wave, and country cutoff date – ESPAD data

Table B.17: Cutoff dates, and quantity of observations per country per wave, for the ESPAD dataset, with indication about the additional presence in the HBSC dataset.

Country	Cutoff date	Wave 1995	2007	2011	2015	All waves	Also in HBSC
Albania	Jan 1st	0	0	3,165	2,519	5,684	
Austria	Sep 1st	0	2,559	0	3,656	6,215	x
Belgium, Flanders	Jan 1st	0	1,888	1,798	1,771	5,457	x
Bosnia and Herz., Fed. of Bosnia and Herz.	Apr 1st	0	2,973	4,489	0	7,462	
Bosnia and Herz., Serb Republic	Apr 1st	0	2,609	3,074	0	5,683	
Bulgaria	Jan 1st	0	2,323	0	0	2,323	x
Cyprus	Sep 1st	0	0	4,071	2,067	6,138	
Czechia	Sep 1st	0	3,901	3,864	2,710	10,475	x
Denmark	Jan 1st	2,209	867	0	0	3,076	x
Estonia	Oct 1st	0	2,314	2,402	0	4,716	x
Faroes Islands	Jan 1st	471	0	0	0	471	
Finland	Jan 1st	2,148	0	0	0	2,148	x
France	Jan 1st	0	2,880	2,529	2,598	8,007	x
Germany, Bavaria	Oct 1st	0	811	723	858	2,392	
Greece	Jan 1st	0	3,030	5,841	3,169	12,040	x
Ireland	Jan 1st	0	0	0	1,437	1,437	x
Italy	Jan 1st	1,409	0	0	0	1,409	x
Kosovo	Jan 1st	0	0	2,312	0	2,312	
Latvia	Jan 1st	0	2,236	2,592	0	4,828	x
Liechtenstein	Jul 1st	0	0	0	315	315	
Lithuania	Jan 1st	0	2,397	2,474	2,536	7,407	x
Moldova	Nov 1st	0	3,127	1,963	2,554	7,644	
Netherlands	Oct 1st	0	2,091	2,044	1,684	5,819	x
Norway	Jan 1st	3,807	0	0	0	3,807	x
Poland	Jul 1st	7,278	0	0	0	7,278	x
Russian Fed., Moscow	Sep 1st	0	1,963	1,718	0	3,681	
Slovakia	Sep 1st	2,287	2,443	1,949	2,169	8,848	x
Slovenia	Jan 1st	2,420	3,044	3,168	3,473	12,105	x
Sweden	Jan 1st	3,467	0	0	0	3,467	x
Ukraine	Jan 1st	6,365	2,411	2,210	2,259	13,245	x
All countries		31,861	45,867	52,386	35,775	165,889	

Note: Serb Republic is one of the two entities of Bosnia and Herzegovina, it is not a different name for Serbia.

Table B.18: Perceptions about risky behaviors.

Variable	N	Mean
Perceived risk of harming themselves		
Occasional cigarette	153,744	0.412
4-5 drinks a week	152,196	0.894
Occasional marijuana	144,503	0.756
Try ecstasy	133,925	0.751
Try amphetamine	135,651	0.944
Perceived peers' consumption prevalence		
Smoke	129,093	0.700
Drink	128,732	0.783
Drunk	124,021	0.478
Marijuana	128,113	0.147
Tranquilizers	126,078	0.033
Ecstasy	128,315	0.031
Inhaler	127,383	0.035

Note: All these outcomes on perceptions about risky behaviors are dummy variables.

Table B.19: Descriptive statistics for expected relative age, expected absolute age, control variables.

Variable	N	Mean	SD
Expected relative age (years)	165,889	0.498	0.305
Expected absolute age (years)	165,889	15.969	0.408
Female	165,889	0.514	
Parents	163,828	0.754	
SES: Low	131,868	0.099	
SES: Medium	131,868	0.512	
SES: High	131,868	0.389	

Note: SES: Low is the reference dummy for family SES. Female, Parents, and the SES variables are dichotomous variables; thus, standard deviation (SD) is not reported. Analyses additionally include vectors for wave, country and season of birth.