Bullied because younger than my mates?
The effect of age rank on victimization at school*

Rosario Maria Ballatore† Marco Paccagnella‡ Marco Tonello§

This version:May 29, 2017

Abstract

Using census data on three cohorts of 5^{th} grade Italian students we investigate how the ordinal rank in the within-school age distribution affects the probability of being bullied. Identification is achieved by exploiting within-school variation in the age composition of different cohorts, and through an IV strategy based on the discontinuity in the probability of enrolling in a given school year generated by an end-of-year cut-off rule. We find that being in the upper part of the school age distribution reduces the probability of being bullied: a one-decile increase in the within-school rank decreases the probability of being victimized by about one percentage point. The effects are stronger for females, children from disadvantaged backgrounds, and children spending the entire day at school; they do not depend on the choice of the reference group, as defined according to socio-demographic characteristics.

JEL Classification: I21, J24, Z13

Keywords: bullying, school violence, ordinal rank, relative age

*We owe special gratitude to Gordon Dahl for discussing a previous version of this paper during a workshop at the Bank of Italy. We also thank, withouth implications, Francesco Avvisati, Francesca Borgonovi, Lorenzo Cappellari, Michela Carlana, Marta De Philippis, Claudio Lucifora, Federica Origo, Judit Pál, Michele Pellizzari, Paolo Sestito, Giovanni Soggia, and seminar participants at the London School of Economics, the Bank of Italy, Invalsi, the 2016 European Association of Labour Economists Conference, the 2016 Annual Meeting of the Italian Association of Labour Economists, and the 2015 Brucchi Luchino Labour Economics Workshop, for useful comments and suggestions. We are indebted to Patrizia Falzetti and Valeria Tortora (Invalsi) for generous support in providing the data. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Italy, nor of the OECD or of its member countries. The usual disclaimers apply.

†Bank of Italy, Regional Economic Research Unit, Cagliari. E-mail: rosariomaria.ballatore@bancaditalia.it.

‡OECD, Directorate for Education and Skills, Paris & Bank of Italy. E-mail: marco.paccagnella@oecd.org (corresponding author).

§Bank of Italy, Regional Economic Research Unit, Florence and CRELI-Catholic University, Milan. E-mail: marco.tonello@bancaditalia.it.

1 Introduction

School violence, as well as more general forms of misbehavior, are a widespread phenomenon that is increasingly capturing the attention of policy-makers, administrators, teachers, and parents. In 2015 20% of 15-year old students reported being victim of some form of bullying on a monthly basis, and 50% reported having being victimized at least a few times per year (OECD 2017).[1] In the US, in the years between 2005 and 2013, about one third of the students aged between 12 and 18 interviewed in the *School Crime Supplement* of the *National Crime Victimization Survey* reported having suffered from some form of bullying at school during the school year, such as being insulted, threatened, stolen things, hurt, or forced to do something against their will (Robers, Zhang, Morgan, and Musu-Gillette 2015). Our empirical analysis exploits data on Italy, where about 50% of adolescents aged between 11 and 17 reported having been victim of some form of violent or disrespectful behavior (ISTAT 2015).

Bullying intrinsically involves a relationship between two actors, with the victim being usually in a weaker position, in terms of age, social background, or physical or psychological development. Typically, bullies beat their victims, steal personal belongings, or force them to do something without consent (Olweus 2013). Being victimized has both short- and long-term consequences. Amongst the former, the psychological literature has traditionally looked at outcomes such as anxiety, insecurity, anti-social behavior, self-esteem, and absences from school.[2] The latter may include lower levels of educational attainment, which then possibly brings about worst labor market outcomes.

Despite its relevance in the process of human capital formation and the potentially damaging effects on subsequent labor market outcomes, school bullying has started to receive attention from the economic literature only recently, thanks to both an increasing awareness of the seriousness of the problem, and to the availability of suitable micro-level data.

Recent attempts to identify a causal link from victimization to such outcomes include Brown and Taylor (2008), Ammermüller (2012), Ponzo (2013), and Eriksen, Nielsen, and Simonsen (2014). On the other hand, convincing empirical evidence on the determinants of being a victim of bullying is still relatively scarce.

In this paper we investigate how age affects the probability of being bullied at school. Age is an important variable to look at, especially from a policy perspective. It is easily

[1] Exposure to bullying at least a few times per year ranges from 27% in Korea to 69% in Latvia. Victimization at least a few times a month involves as few as 9% of students in the Netherlands and as many as 26% of students in New Zealand (OECD 2017).

[2] See Juvonen and Graham (2014) and Olweus (2013) for recent reviews.

observable, and school principals can take it into account when deciding on school admission or class formation.

Age could affect the probability of being bullied at school in two main different ways. First, younger children could be less prepared for school just because of lower physical or psychological development, determining differences in outcomes such as school performance, self esteem and also bullying. This is the effect that comes with being younger in an *absolute* term: it is simply related to children development *per se* and does not depend on the age composition of the peers a student is exposed to (henceforth, we label this component "absolute age effect").[3] Second, age could affect the probability of victimization also by determining the student's relative position with respect to her schoolmates (henceforth, "relative age effect"). Depending on the age composition of a particular school-cohort, two students of the same (absolute) age may well result among the oldest in the age distribution of one cohort and among the youngest in that of another cohort. As bullying involves a relationship between at least two actors, we argue that this latter dimension is the one that matters more in explaining the probability of victimization. For this reason, our work focuses on the causal relationship between relative age and victimization, and we use absolute age at school to control for differences in physical and psychological development that are correlated with both relative age and with the likelihood of being bullied.

Traditionally, the literature has measured relative age as the linear distance from the average age of the peer group (Cascio and Schanzenbach 2016; Elder and Lubotsky 2009; Mühlenweg 2010). In this paper we use instead a cardinal measure of relative age, namely the ordinal rank within the school-cohort (i.e. the *age rank*). We believe that there are at least two advantages in using this measure. First, ordinal rank allows for more informative comparisons between students attending school cohorts with very different age distributions. Students that are 3 months older than the average age of their schoolmates could very well be at different percentiles of the school-cohort age distribution. Cardinal measures of relative age are not able to capture such hierarchical differences, that have been proved to be important in determining behavioral outcomes such as leadership (Dhuey and Lipscomb 2008) and risky behaviors (Elsner and Isphording 2017b). Second, the use of age rank allows us to set up a more flexible estimation framework, as identification can be achieved exploiting also higher moments of the age distribution (e.g. variance, skewness and so on). For instance, by

[3] Cascio and Schanzenbach (2016) formalized the breakdown of the absolute age effect in three components. First, the school starting age effect (SSA) originating from starting school earlier or later. The second component depends from age at the time the test is taken (i.e. the Age At Test effect, AAT). Third, there is an effect of the total number of years of schooling (i.e. the Years Of Schooling effect, YOS). These three components are linearly dependent ($AAT = SSA + YOS$) making it impossible to separately estimate the contribution of each single effect.

including school-by-cohort fixed effects we can compare students that have the same absolute age and the same distance from the mean, but result in different ordinal positions because the shape of age distribution differs.

Our work provides three main contributions. First, we add to the existing literature on school bullying by focusing on one potential determinant (i.e. age rank) that has never been looked at before. In this regard, our paper expands on the work of Mühlenweg (2010), who focuses on the link between relative age and school victimization. However, differently from the estimates provided in that paper, which incorporates both the absolute and the relative age effects, we are able to provide separate estimates of these two parameters. Second, our work contributes to a broader literature on the effects of age on behavioral and non-cognitive students' outcomes (Crafword, Dearden, and Greaves 2015; Dee and Sievertsen 2015; Lubotsky and Kaestner 2016; Pellizzari and Billari 2012). We improve on existing works by looking at an outcome which has not yet been taken into account (i.e. school victimization), and, again, by focusing for the first time on a different dimension of relative age, i.e. the ordinal rank in the age distribution. To the best of our knowledge, this is the first work that specifically focuses on the age rank effect, thus providing the literature with a new parameter of interest. Finally, we add to a new but rapidly growing literature that looks at the effect of *achievement* rank in education (Elsner and Isphording 2017a,b; Murphy and Weinhardt 2014; Tincani 2017) by showing that, in addition to academic achievement, also the rank within the age school-cohort distribution has important effects on behavioral outcomes.

Our data cover the entire population of Italian students enrolled in 5^{th} grade in three different school years. Identification relies on cohort-to-cohort variation in the within-school age distribution (Elsner and Isphording 2017a,b; Murphy and Weinhardt 2014).[4] We also employ an IV strategy, instrumenting age with the theoretical age the children should have had at the time of school enrollment. This is a standard practice in the literature on age effects (Black, Devereux, and Salvanes 2013). In Italy the school year starts in September, and children are expected to enroll in the first grade of primary school the calendar year in which they turn six. December 31^{st} thus generates a discontinuity in children' age at the time they start school. Such discontinuity, however, is not sharp, because children born between January and April are allowed to enroll in the year they turn five. Our IV strategy therefore

[4]In our data we observe a school panel composed by repeated cross-sections of students enrolled in 5^{th} grade. Therefore, when we refer to *cohorts* we always mean *school-entry cohorts*, unless otherwise specified. In our setting school-entry cohorts essentially coincide with survey waves, as grade repetition is negligible in primary schools. They do not coincide with *year-of-birth cohorts* because of the possibility of anticipating or postponing school entry that we discuss in detail in a later section.

addresses a major threat to identification, i.e. the strategic behavior of parents that, by choosing the timing of school entry, determine the assignment of their children to a given school-cohort.

We find that a student's ordinal rank significantly affects the likelihood of being victimized at school. In our preferred specification, a one-decile increase in the ordinal rank decreases the probability of being victimized by 0.9 percentage points, against a baseline probability of about 40%. The result is robust to several alternative empirical specifications. An investigation of the effects across and within different subgroups also reveals interesting patterns. When we estimate separate effects according to student's gender, socio-economic and immigrant status, and time spent at school (half- or full-day), we find larger effects for girls, for immigrants, for children coming from low socio-economic backgrounds, and for children who spend the full day at school. When instead we change the reference group for which we compute the student's ordinal rank (looking, for instance, at the rank of a male student within the age distribution of males), we find much smaller effects, that never reach common thresholds for statistical significance, suggesting that the relevant reference group (i.e. the group within which the age rak affects victimization) is not dictated by observable characteristics, but rather refer to the entire school pool of peers.

The rest of the paper is organized as follows. Section 2 describes the institutional setting and the data used. Section 3 presents the empirical strategy, while Sections 4 and 5 discuss the main results, present some robustness checks, and investigate relevant heterogeneities in the estimated effects. Section 6 concludes.

2 Institutional setting and data

2.1 The Italian school system and the SNV surveys

In Italy, compulsory schooling starts with five years of primary school (grades 1 to 5, corresponding to ISCED level 1), and then continues with three years of junior high school (grades 6 to 8, ISCED level 2). Children normally enroll in the first grade of primary school the year they turn six. Up to grade 8, the school curriculum is identical for all students.

Starting from the 2009/10 school year, the National Institute for the Evaluation of the Education System (Invalsi) has carried out annual evaluations of students' achievement by means of the National Students' Assessment Survey (SNV, from the Italian acronym). The SNV includes a direct assessment of mathematics and reading, and a background questionnaire (so called *Student Questionnaire*). The SNV takes the form of an annual census, since it is compulsory for all students attending second, fifth and sixth grade. For each grade, about 400,000 students sit the assessment every year, over two different days (for the two subjects) in late May.

SNV data contain therefore information on test scores, individual and family background characteristics, collected from three main sources: (*i*) test scores data are collected through the standardized assessments in the two subjects; (*ii*) students' characteristics are directly compiled by the school administrative staff, drawing from school administrative records, and appear on each student's answer sheet; (*iii*) additional individual-level information on family, school and environmental characteristics are collected through the *Student Questionnaire*, which is taken by 5^{th} grade students the same day of one of the two tests (after finishing the test). Invalsi also combines information on parental background (in particular mother's and father's occupation and educational attainment) with information on home possessions (such as the presence of a computer and of Internet connection and the number of books at home) to construct a standardized individual-level index of Socio-Economic Status (SES).

The data include anonymous school identifiers, which makes it possible to follow schools over time. This is crucial to implement our empirical strategy of exploiting within-school cross-cohorts variations in the age composition of schoolmates. Moreover, using the universe of the students enrolled in a given grade entails a notable advantage for our identification strategy, as it allows to precisely identify the entire group of relevant peers, overcoming possible problems of attenuation bias (Micklewright, Schnepf, and Silva 2012). Crucially for our purposes, the *Student Questionnaire* for 5^{th} graders contains a set of questions on students' behavior at school and outside school. From these questions, we retrieve the information on students' victimization. Finally, we were able to gain access to the month of birth of each student, which we use to compute students' age.

2.2 Data and descriptive evidence

We exploit three waves of the SNV survey, covering the school years 2010/11, 2011/12 and 2013/14, and we focus on 5^{th} grade students, for which we can retrieve information on victimization from the *Student Questionnaire*.[5] From the census of primary schools, we drop those in which less than 10 students are enrolled in 5^{th} grade, as well as schools that are observed in one wave only (less than 2% of the schools). Finally, we exclude all individuals who do not provide answer to the victimization question on the *Student Questionnaire* (about

[5]Unfortunately, we cannot exploit the 2009/10 wave because it does not contain a school identifier which can be linked to the other waves. We cannot use the 2012/13 wave either as in that year no questions on bullying were present in the *Student Questionnaire*.

4% of the overall sample of students).

[Table 1 about here]

Panel A of Table 1 presents some descriptive statistics based on our final sample of more than 1,200,000 students and 14,300 schools. In line with existing works that use data from international surveys (Mühlenweg 2010; Ponzo 2013), we identify students as victims of bullying whenever they answered positively to any of the following questions: (*i*) the student was beaten; (*ii*) the student was forced to do something against her will; (*iii*) the student was stolen things. About 40% of students in our sample declare having been victim of some sort of bullying at school. The probability of being victimized is higher for males, immigrants, students coming from low socio-economic backgrounds, and students spending the entire day in school (Table 1, Panel B).

These figures are in line with other data sources and with data on other countries. Across OECD countries participating in PISA in 2015, 49% of students reported being victimized at least a few times per year (OECD 2017). In 2003, between 24% and 47% of 4^{th} grade students participating in the Trend in International Mathematics and Science Study (TIMMS) declared having been hurt or hit by a schoolmate in the month before the survey (35% in Italy), while 12% to 32% reported having been stolen things (30% in Italy; see Ammermüller 2012).

We use information on year and month of birth to derive students' age. The average student in our sample is about 10-years-old (121.6 months).

To compute the age rank, we follow Elsner and Isphording (2017a), converting the absolute ordinal rank into a percentile rank.[6] After assigning a value of 1 to the lowest-ranked, and N to the highest-ranked student in a given school in a given cohort (where N is the size of the school cohort), the percentile ranks are computed as:

$$\text{percentile rank} = \frac{\text{absolute rank} - 1}{N - 1} \tag{1}$$

Figure 1 plots the main features of the distribution of the age rank (as calculated from Equation 1 and netted out of school and cohort fixed effects, vertical axis), against the deciles of the age distribution in the whole sample (horizontal axis). The figure thus illustrates the type of variation we exploit, by showing that, for a given decile in the overall, entry-cohort-wide age distribution, there is quite a bit of variation in the *age rank* measure, calculated within a specific school. Similarly to Elsner and Isphording (2017a), we can also show that

[6] The standardization is necessary to compare ordinal rank measures referred to peer groups of different size.

our data display substantial within school variability in the ordinal rank measure: for any given age, the within-school standard deviation of the age rank, conditional on school and cohort fixed effects, amounts to 0.122, meaning that students with the same age may end up in different rank positions depending on the school-cohort they belong to.[7]

[Figure 1 about here]

Notice that the construction of age rank is based on the school, and not the class dimension. Two reasons lie behind this choice. First, data from victimization surveys show that the majority the interactions that involve episodes of bullying occurs outside the classroom, for instance during the recreation time in the hallway or on the school courtyard, in bathrooms or locker rooms, in the cafeteria, or on the school bus (Robers et al. 2015). Second, the process of class formation makes it very difficult to exploit exogenous variability in the age rank when computed at class level. Classes are formed in an almost arbitrary way by school staff, and while there are broad guidelines to ensure a balance composition of classes (e.g. along the gender dimension), it is rather common knowledge that selection occurs among a number of unobserved dimensions, the most common probably being parents lobbying to have their children assigned to the class taught by the best teachers, or not to be separated from past friends.

3 Empirical strategy

3.1 The ideal experiment

As our goal is to investigate how ordinal rank in the age distribution affects the probability of being victimized at school, we start from the following baseline equation:

$$b_{isc} = \alpha + \beta r_{isc} + \gamma X_{isc} + \epsilon_{isc} \tag{2}$$

where b_{isc} is a dummy variable indicating whether student i in school s in cohort c reports to having been victimized, and r is the age rank of the student in the school-cohort age distribution, as defined in the previous section. X_{isc} is a vector of student-level characteristics including age (in months), the index of Socio-Economic Status (SES), dummies for first- and second-generation immigrants, and a gender dummy; ϵ_{isc} is the error term, which we allow to be correlated within schools.[8]

[7] Considering the mean sample school cohort size of about 50 students, for a given age, the age rank would vary by almost 5 rank positions.

[8] As argued by Cameron and Miller (2015), the school level is the appropriate level at which standard errors should be clustered, given that we have a short panel component (i.e. three school years).

In an ideal experiment, the researcher would like to vary the ordinal rank of students in the school-cohort age distribution through random assignment to different peer groups (e.g. different schools or different cohorts within the same school). This is illustrated in the Appendix Figure A1. Exogenous variation in the ordinal rank comes from the assignment of students to different peer groups, that can be characterized by a different average age (as in peer group B, whose age distribution is simply shifted with respect to peer group A), or by a different dispersion (as in peer group C, that is characterized by a more compressed age distribution, but whose average age is the same as group A). If students were randomly allocated to different peer groups, it would be enough to estimate Equation 2 by OLS to recover the causal effect of the age rank on the probability of being victimized, as students' assignment to a given peer group would be entirely random. The assumption of random assignment is difficult to defend in our setting, as parents can choose not only the school, but also - following the administrative rules detailed above - the school year in which to first enroll their children, and therefore the school-cohort.

thus, there are two main reasons behind the possible endogeneity of the age rank measure. First, there can be selection into different *schools*, characterized by idiosyncratic and unobserved differences (not only in terms of the age distribution, but also in terms of a variety of factors, like teaching quality, reputation, school climate, and so on). Second, there can be selection into different *cohorts*, as parents have some degree of autonomy in choosing the academic year in which they enroll their children in primary school, and this might alter - in a selected way - the ordinal rank within each school cohort. In what follows, we illustrate how we tackle these two threats to identification.

3.2 Selection into schools

In order to take care of selection into schools, we exploit the longitudinal nature of our data and include in Equation 2 a full set of school-by-cohort fixed effects ($\delta_c \times \lambda_s$):

$$b_{isc} = \alpha + \beta r_{isc} + \gamma X_{isc} + \delta_c \times \lambda_s + \epsilon_{isc} \quad (3)$$

The estimation is therefore based on idiosyncratic variations in the age composition of different cohorts of students that enter the same school in different years. While a specification with school-by-cohort fixed effects is essentially equivalent to a within-transformation of all variables at the school-cohort level, the coefficient on age rank is still identified from differences in the variance and in higher moments of the age distribution, as illustrated in Figure A1 (i.e. comparing peer group A with C). At the same time, this specification makes it possible to estimate the effect of age rank holding constant individual age and the average age of schoolmates, thus controlling for a cardinal measure of relative age as usually defined in the literature (Cascio and Schanzenbach 2016; Elder and Lubotsky 2009).

Furthermore, school-by-cohort fixed effects control not only for selection into schools, but also for any unobserved mean differences between different school cohorts, as well as for any unobservable group shock. In particular, they take account of dynamic selection into schools (i.e. the fact that schools can become over time more or less attractive to a particular group of students), and, more generally, of any differences in cohort composition or in school inputs. In the Italian context, for instance, it is particularly important to control for differences in teacher quality: given that primary school teachers tend to teach the same cohort of students, following them from the first to the fifth grade, different cohorts of students are normally taught by different teachers.[9]

3.3 Selection into cohorts

Even after the inclusion of school-by-cohort fixed effects, identification still relies on the assumption that belonging to a given cohort is as good as random. This is difficult to defend in our setting, as parents can manipulate, within specific rules, the timing of school entry. In Italy children are normally required to start primary school in the year in which they turn six (see section 2): December 31^{st} therefore generates a natural cut-off date that changes, in a discontinuous way, both the relative position of a student within the school and the observed age at the time of the survey. The rule is flexible to some extent, as children born between January and April are allowed to enroll earlier, i.e. in the year they turn five (this is what we label *early enrollment*).[10] Parents are also allowed to postpone children enrollment (so-called *late enrollment*), but this latter case is much less frequent (as illustrated in Figure 2), and it is usually motivated by specific problems such as disabilities or, in the case of non-native students, insufficient proficiency in the language of instruction.

[Figure 2 about here]

The enrollment rule generates a theoretical starting age function (Equation 4), which is a discontinuous function in the month of birth:

$$f_{isc} = 72 - (9 - m_{isc}) \tag{4}$$

[9] Alternative specifications will be discussed in a later section. In our setting, teacher quality can be thought of in broader terms, as having an impact not only on test scores, but also on students' behavior and discipline, thus influencing the probability that episodes of bullying take place.

[10] The administrative rules are detailed in the Ministry of Education Regulations (*Circolare Ministeriale No. 4/2009*).

[Figure 3 about here]

Figure 3 displays the relationship between the theoretical starting age function (as expressed by Equation 4) and the school starting age observed in the data: compliance with the theoretical enrollment rule based on the 31^{st} December cut-off is good, although not perfect, especially for children born in the first months of each year, whose parents clearly opted often for an early enrollment. Late and early enrollment can be manipulated by parents and by the school staff: parents can request early enrollment for their children, but such request can very well not be satisfied, as the final decision on admission rests in this case with the school principal. Clearly, both the requests from the parents and the final decision of the school principal are based on the observation of several characteristics of the child that are unknown to the researcher and that are plausibly correlated with the outcomes of interest. The fact that early and late enrollees are a selected population is also illustrated in Table 2, which shows that, for instance, females and children with a higher socio-economic status are more likely to enroll earlier, while immigrants, especially of first-generation, are generally more likely to postpone enrollment.

[Table 2 about here]

To tackle the endogenous sorting into cohorts, which contaminates the estimated coefficient of both age and age rank (which is a function of age), we exploit the 31^{st} December cut-off as a source of exogenous variation, and, given that compliance is not perfect, we instrument observed age and observed ordinal rank with their theoretical counterparts. That is, the age and age rank variable are instrumented, respectively, with f_{isc} and with the theoretical percentile rank constructed starting from the theoretical starting age (f_{isc}).

The identification strategy thus exploits an IV design, where the exogenous variation in the instrumental variable is given by the cut-off rule for enrollment in the first grade of primary school in the year the child turn six (Black et al. 2013). A potential threat to this identification strategy is constituted by endogenous selection into seasons of birth, due to parental fertility choices (see Carlsson, Dahl, Öckert, and Rooth 2015 and Black et al. 2013, among others). To address this issue, in section 4.2 we show that our results do not vary if we restrict the estimation sample to students born in months close to the cut-off date. Finally, it is worth noting that the IV strategy described so far also helps in addressing possible bias from measurement error. As we only have access to the month of birth, our age variables are all measured with error, which would induce downward bias in the estimated parameters.

4 Results

4.1 Baseline results

We estimate linear probability models of several variants of Equation 3, with robust standard errors clustered at the school level.[11] The first three columns of Table 3 report results from OLS regressions, while columns 4 to 6 report results that make use of instrumental variables. All estimates include school-by-cohort fixed effects; columns 2 and 5 also include individual level controls, while in columns 3 and 6 we add year of birth fixed effects to take into account potential unobserved heterogeneity at the birth-cohort level. In all specifications, *age rank* is estimated to have a negative and statistically significant effect on the probability of being victimized. OLS estimates indicate that a one-decile increase in the age distribution is associated to a decrease in the probability of being bullied between 0.3 and 0.7 percentage points.

[Table 3 about here]

However, as illustrated in the previous section, OLS estimates could be biased as both age rank and age could be correlated to unobserved variables contained in the error term. In the Italian setting, in particular, we have shown that early school enrollment is a rather common phenomenon. Early enrollment is an explicit choice made by the parents, likely based on information that is unobservable to the econometrician.[12] Early enrollees are thus likely to be a selected group of students under multiple dimensions. They are generally more psychologically mature, and if this maturity were associated to a lower propensity to be victim of bullying, then OLS estimates would be upward biased, as the sample would contain students that, while being at the bottom of the age distribution, have lower chances of being victimized. Furthermore, the regressor of interest is measured with error, as we are only able to observe the month of birth, and not the exact day of birth. This induce an attenuation bias in the OLS specifications, pushing the estimated coefficients toward zero. Results from columns 4 to 6 of Table 3 are consistent with these considerations. The coefficients estimated through an IV strategy are always larger (in absolute terms) than the corresponding OLS coefficients. Focusing on our baseline IV specification, which includes the full set of fixed effects and individual controls (column 6), we find that increasing the age rank by one decile would decrease the probability of being victimized by almost one percentage point, against

[11] Linear probability models can be more flexibly used in a 2SLS regression framework as compared to probit or logit models. The OLS results, however, do not change substantially if we adopt either a logit or a probit specification.

[12] Kindergarten teachers, for instance, usually give advice to parents on the school readiness of each child.

a baseline probability of 40%. This would mean that, for instance, moving a student from the top to the bottom of the age distribution would increase the chances of victimization by almost one quarter.

Our results also show that age *per se* does not seem to be a determinant of being bullied at school, once accounting for its endogeneity in the IV model (see the extended results reported in the Appendix Table A1). In his work on the academic consequences of being bullied at school, Ammermüller (2012) also provides descriptive evidence of the individual level correlates of school victimization based on a student fixed effects model. His mixed findings on the age effects (null or negative, depending on the specification) are somehow in line with what we find. However, he does not focus on rank effects, nor on the distinction between the absolute and the relative age components. Mühlenweg (2010), who takes into account the endogeneity of students' age using an IV strategy similar to ours, shows that the probability to suffer from school victimization is reduced by 8 percentage points for older children. However, such estimate incorporates both the effect of being older in an absolute sense and as compared to school peers, and does not focus on the age rank distribution.

4.2 Robustness

To test the robustness of our results, Table 4 reports various alternative specifications. Threats to identification would arise in case parents could manipulate the season of birth of their children. To address this potential concern, in column 1 we restrict the sample to students born around the 31^{st} December threshold (more precisely, in the first and fourth quarter of the year), with the idea that, within that restricted window, the assumption of random birth on either side of the threshold is more tenable. In this specification, the estimated coefficient is larger than in our baseline specification: a one-decile increase in the ordinal rank would decrease the probability of being victimized by almost 2.7 percentage points. The stronger effect that we find in the discontinuity sample is justified by the fact that we are focusing on the subsample of students who might opt for early enrollment, for whom the LATE effect estimated through 2SLS is stronger, as compared to the same effect on the full sample of students.

[Table 4 about here]

Another potential threat to identification comes from the fact that the school-by-cohort fixed effects estimation strategy (even when we also account for year of birth fixed effects) does not take into account relevant age variations at the school-by-year of birth level. Specifications in columns 2, 3 and 4 show that the baseline estimates do not change significantly when we control for the mean, the variance, or both mean and variance of the age variable, calculated at the school-by-year of birth level. Specifications in columns 5 and 6 include higher polynomials of the age variable, instrumented for with the corresponding higher moments theoretical counterparts. This exercise shows that the results do not depend on the way in which the age variable enters in the vector of controls, whether linearly or not.

An alternative estimation strategy could make use of separate school and cohort fixed effects, in a two-way fixed effects model (Elsner and Isphording 2017a). This specification would have the advantage of being less computationally demanding as compared to the school-by-cohort fixed effects model. However, such model would be unable to take into account the school-by-cohort unobserved factors illustrated in section 3.2. To allow for comparability across the two models, column 7 shows the results of the two-way fixed effects model, in which we also add as control variables the school-by-cohort average share of females, share of immigrant students, and SES index. It is noteworthy that while the school-by-cohort fixed effects model estimated so far ruled out comparisons between peer groups A and B of figure A1, estimates of column 7 exploit variability in the rank position that comes even from differences in the mean of the school cohort age distribution. The estimated coefficient of the age rank is still close to our baseline results, although less precisely estimated, meaning that school-by-cohort confounding effects are important to control for as they add precision to the estimates, but not strong enough to alter significantly the ordinal rank effect.

5 Heterogeneous effects

In this Section we investigate whether the age rank effect varies across individuals sharing the same observable characteristics, and whether it is sensitive to the choice of the reference group according to which the age rank is computed. In a sense, the former exercise looks at heterogeneity *between* groups, while the latter is a search for effects *within* groups.

5.1 Between-groups effects

Knowledge of which are the types of individuals that, based on simple observable characteristics, are more 'at risk'of being victimized is obviously of interest to school staff, like teachers and principals, and to policy makers in order to better target *ad hoc* policies to combat bullying at school. We are able to look at groups defined according to gender, immigrant status, socio-economic status (above or below the median value of the SES index), and amount of time spent at school (half- or full-day time schedule).[13]

[Table 5 about here]

The results in Table 5 show that the age rank effect is slightly more pronounced for females, for immigrant students (although the estimate is much less precise and does not reach conventional levels of statistical significance, plausibly due to the limited number of immigrants in our sample), for students coming from a lower socio-economic background, and for students spending longer time at school. With the exception of this last group, for which the effects are almost twice as large than our baseline, we note that estimated coefficients do not vary much, and are reasonably close to the baseline.

5.2 Which is the relevant peer group?

The previous subsection looked at *between-groups* heterogeneity, which is informative on whether the effect of ordinal rank is different for different groups of students. Another aspect which is interesting to analyze entails a focus on *within-groups*: looking at the ordinal rank in the within-group age distribution, rather than in the age distribution within the entire cohort of schoolmates, might provide hints about what are the relevant peer groups for the playing out of age rank effects. This ultimately corresponds to understanding whether students identify peer groups based on gender, immigrant or socio-economic status, time spent at school, and thus whether the age rank affects the probability of victimization in a different way within different groups. To this end, we simply retrieve the ordinal rank measure restricting each time the calculation to the subsamples defined above.

[Table 6 about here]

The results of such analysis are presented in Table 6. All estimated coefficients in the within-group specifications are much smaller than the baseline, and are in no case statistically significant. This suggests that students do not sort out according to the socio-demographic characteristics analyzed here. Rather, the relevant peer group (for the purpose of the age rank effect on school victimization) is constituted by the entire population of schoolmates.[14] What

[13]Parents can choose between a schedule with 30 or 40 hours of lessons per week. Such requests are accommodated conditional on the availability of sufficient resources at the level of individual schools. On the part of families, the request for longer schedules are usually dictated by the need to conciliate work and childcare commitments. In schedules within the amount of 30 hours students stay at school in the morning only, while schedules from 31 to 40 hours entail a variable number of afternoons spent at school.

[14]It would be very interesting to estimate the effect within the group of classmates, rather than schoolmates. However, we are not able to follow classes over time, which prevents us to include class-by-cohort fixed effects. As the allocation of students into classes is certainly non random, we believe we are not in a position to provide convincing causal evidence on the age rank effect at the class level. Furthermore, descriptive evidence in the

seems to matter the most is the amount of time spent at school (although even this does not reach conventional levels of statistical significance), which is likely to increase the frequency of interactions and therefore to strengthen ties (positive or negative) among schoolmates.

6 Concluding remarks

Being bullied at school is a growing phenomenon that has only recently grasped the attention of policy makers. Being bullied at school has been found to have both short and long run negative effects on cognitive and non-cognitive outcomes, but also on labor market performance and health (Bowes, Joinson, Wolke, and Lewis 2015). Recent studies make considerable contributions to this growing body of the literature by adopting more sophisticated identification strategies (Ammermüller 2012; Eriksen et al. 2014; Ponzo 2013), which tend to confirm the short-term negative effects for cognitive outcomes. Yet, convincing evidence on the determinants of school victimization is still lacking. Several studies in psychology and other social sciences have shown that bullying is associated to many individual (observable and unobservable) traits such as physical appearance, race, sexual preferences, low self-esteem, depression, loneliness (Juvonen and Graham 2014; Olweus 2013). However, it is often not clear from these works which is the direction of the causality link.

Understanding the determinants of school victimization is of crucial importance for the design of policies aimed at countering bullying at school. This paper aims at start filling this gap in the literature by estimating the effects of one potentially important channel, namely age. Age is an important policy lever, as it is easily observable and it is often one of the factor that informs decision about school admission and class composition. We find that the ordinal rank in the within-school age distribution has a significant effect on the probability of being victimized. In particular, a one-decile increase in the distribution reduces the probability of being victimized by about one percentage point (in our preferred specification), against a baseline probability of 40%. This implies that moving a student from the top to the bottom of the distribution would increase the chances of victimization by almost one quarter.

Our results also help to better understand interactions within schools, underlying the importance played by rank or status, as stressed by recent works on peer effects (Elsner and Isphording 2017a,b; Murphy and Weinhardt 2014; Tincani 2017). A practical implication of our results is that, while the age composition of classes and schools certainly plays an

literature suggests that most episodes of bullying occur outside the classroom (for instance at the canteen during lunch-time, or in the schoolyard during recreational intervals, or even outside the school premises at the end of the school day). For these reasons, we believe the school level is still an important dimension to look at.

important role, policies can do very little to limit the effects of ordinal rank. A higher awareness of such effects on the part of teachers and school staff can however help them in identifying more at-risk students and design appropriate targeted interventions.

References

Ammermüller, A. (2012). Violence in European Schools: A Widespread Phenomenon that Matters for Educational Production. *Labour Economics 19*(6), 908–922.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2013). Under Pressure? The Effect of Peers on Outcomes of Young Adults. *Journal of Labor Economics 31*(1), 119 – 153.

Bowes, L., C. Joinson, D. Wolke, and G. Lewis (2015). Peer victimisation during adolescence and its impact on depression in early adulthood: prospective cohort study in the united kingdom. *BMJ 350*.

Brown, S. and K. Taylor (2008). Bullying, education and earnings: Evidence from the national child development study. *Economics of Education Review 27*(4), 387–401.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317 – 372.

Carlsson, M., G. B. Dahl, B. Öckert, and D.-O. Rooth (2015). The Effect of Schooling on Cognitive Skills. *The Review of Economics and Statistics 97*(3), 533–547.

Cascio, E. and D. W. Schanzenbach (2016). First in the Class? Age and the Education Production Function. *Education Finance and Policy 3*(11), 225–250.

Crafword, C., L. Dearden, and E. Greaves (2015). The drivers of month-of-birth differences in chidren's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society A 177*(4), 829–860.

Dee, T. S. and H. H. Sievertsen (2015). The gift of time? School starting age and health. NBER Working Paper 21610, National Bureau of Economic Research, Cambridge, MA.

Dhuey, E. and S. Lipscomb (2008, April). What makes a leader? relative age and high school leadership. *Economics of Education Review 27*(2), 173–183.

Elder, T. E. and D. H. Lubotsky (2009). Kindergarten entrance age and children's achievement: Impacts of state policies, family background, and peers. *Journal of Human Resources 44*(3), 641–683.

Elsner, B. and I. E. Isphording (2017a). A big fish in a small pond: Ability rank and human capital investment. *Journal of Labor Economics 35*(3).

Elsner, B. and I. E. Isphording (2017b). Rank, sex, drugs, and crime. *Journal of Human Resources*.

Eriksen, T. L. M., H. S. Nielsen, and M. Simonsen (2014). Bullying in elementary school. *The Journal of Human Resources 49*(4), 839–871.

ISTAT (2015). Comportamenti offensivi e violenti tra i giovanissimi: Bullismo. In *Multipurpose survey on households: aspects of daily life.*

Juvonen, J. and S. Graham (2014). Bullying in Schools: The Power of Bullies and the Plight of Victims. *Annual Review of Psychology 65*, 159–185.

Lubotsky, D. and R. Kaestner (2016). Do 'Skills Beget Skills'? Evidence on the effect of kindergarten entrance age on the evolution of cognitive and non-cognitive skill gaps in childhood. *Economics of Education Review 53*, 194 – 206.

Micklewright, J., S. V. Schnepf, and P. N. Silva (2012). Peer effects and measurement error: The impact of sampling variation in school survey data (evidence from PISA). *Economics of Education Review 31*(6), 1136 – 1142.

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

Murphy, R. and F. Weinhardt (2014). Top of the class: The importance of ordinal rank. CESifo Working Paper 4815, CESifo Munich.

OECD (2017). *PISA 2015 Results (Volume III): Students' Well-Being.* OECD Publishing.

Olweus, D. (2013). School bullying: development and some important challenges. *Annual Review of Clinical Psychology 9*, 751–780.

Pellizzari, M. and F. Billari (2012). The younger, the better? Age-related differences in academic performance at university. *Journal of Population Economics 25*(2), 697–739.

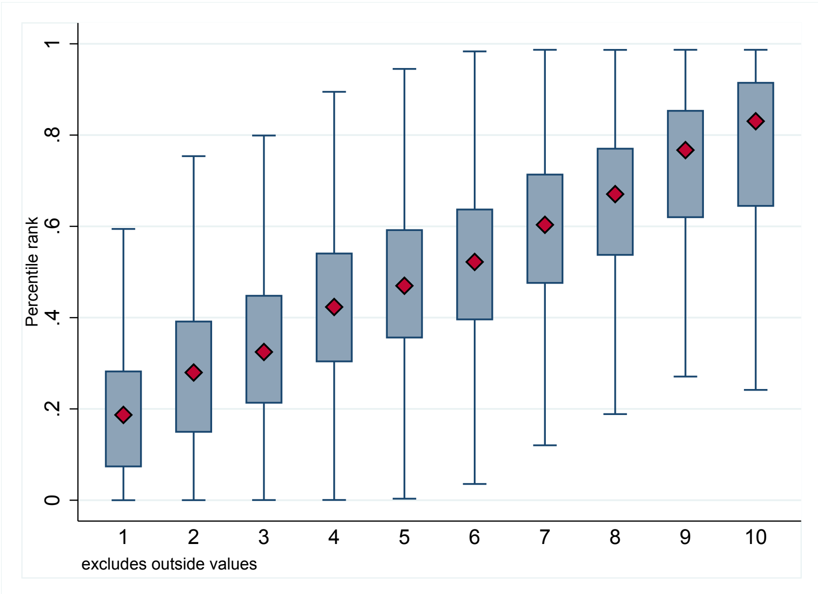
Ponzo, M. (2013). Does Bullying Reduce Educational Achievement? An Evaluation Using Matching Estimators. *Journal of Policy Modeling 35*, 1057–1078.

Robers, S., A. Zhang, R. E. Morgan, and L. Musu-Gillette (2015). Indicators of school crime and safety: 2014. Technical report, National Center for Education Statistics, U.S. Department of Education, and Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice. Washington, DC.

Tincani, M. M. (2017). Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence. CESifo Working Paper Series 6331, CESifo Group Munich.

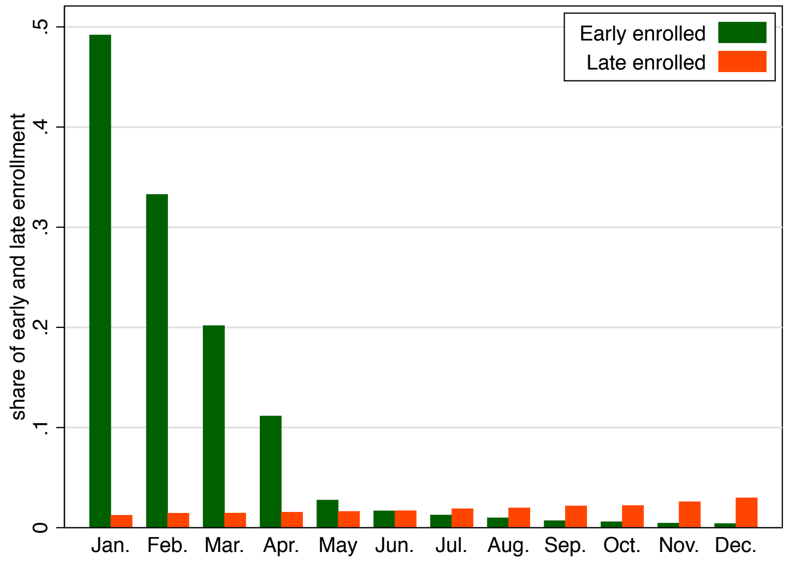
Figures

Figure 1
Global age distribution and local school cohort rank



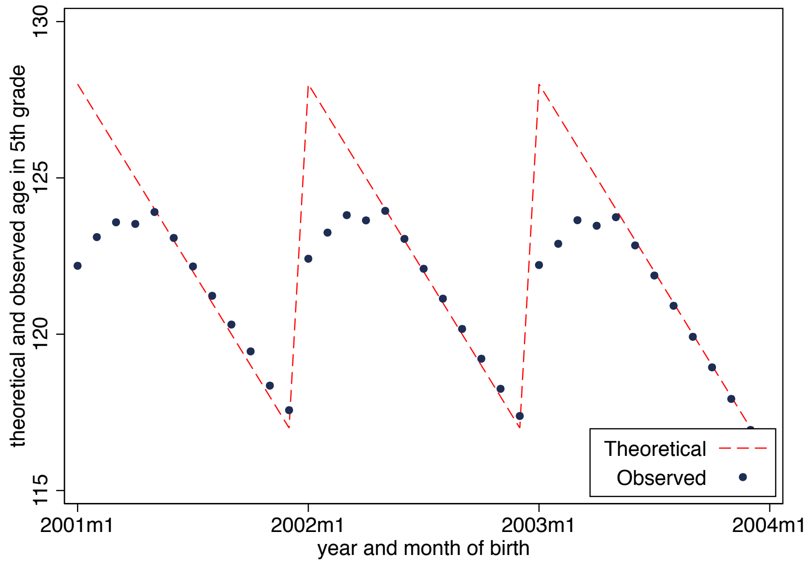
Notes: the graph shows the relationship between the global age distribution and the local rank in a school cohort. The horizontal axis shows the deciles of the global distribution of the age variable (indicated with the numbers 1...10), while on the vertical axis the box plots display the main features of the local rank distribution (the minimum, the maximum, the 25^{th} and 75^{th} percentiles, and the median, indicated with the red diamond). The percentile rank are residuals from OLS regressions on school and cohort fixed effects, and adjusted adding the regression constant so to express them between 0 and 1. Outliers (defined as observations above the 99^{th} and below the 1^{st} percentile) are excluded. *Source*: based on Invalsi SNV (school years: 2010/11, 2011/12, 2013/14).

Figure 2
Early and late enrollments



Notes: the bars show, for each month of birth (horizontal axis), the share of students who enrolled before (*early enrollments*) or after (*late enrollments*) the corrisponding natural school cohort, as defined by the year of birth and the December 31^{st} cut-off rule. *Source*: based on Invalsi SNV.

Figure 3
Theoretical and observed school starting age



Notes: the graph plots the theorethical starting age as defined by the year of birth and the December 31^{st} cut-off rule (dashed line) and the observed starting age (dotted line). *Source*: based on Invalsi SNV.

Tables

Table 1
Descriptive statistics

	mean	sd
Panel A: descriptive statistics		
Dependent variable:		
Bullied (dummy)	0.405	0.491
Control variables:		
Age rank (percentile)	0.499	0.295
Age (months)	121.658	4.526
Socio-economic status Index (SES)	0.106	1.012
Share of females	0.497	0.500
Share of immigrants	0.095	0.294
Share of first-generation immigrants	0.041	0.198
Share of second-generation immigrants	0.055	0.228
Average class size	18.830	3.806
Panel B: probability of being bullied by groups		
Males	0.427	0.495
Females	0.383	0.486
Natives	0.399	0.490
Immigrants	0.463	0.499
High SES	0.391	0.488
Low SES	0.420	0.493
Time at school: half-day	0.391	0.488
Time at school: full-day	0.435	0.496
No. of schools	14,399	
No. of students	1,255,947	

Notes: the Index of Socio-economic status (SES) is provided by Invalsi and it is based on parents' occupation and education level, and on an index of home possessions, capturing the availability of certain goods in the household (such as a computer, an Internet connection, the number of books). The SES Index ranges between -1 and +1, taking positive values for students with socio-economic background higher than each cohort average (which corresponds to 0), and negative values for more disadvantaged students. Immigrants are defined as individuals without Italian citizenship: first generation immigrants are born abroad, while second generation immigrants are born in Italy; a half-day time at school schedule refers to 30 or less than 30 hours of school per week (i.e. students are typically not required to stay at school in the afternoon), while a full-day time at school schedule corresponds to 32 to 45 hours of schooling per week (i.e. students are required to stay at school at least two afternoons). *Source*: based on Invalsi SNV (school years: 2010/11, 2011/12, 2013/14).

Table 2
Students' observable characteristics and timing of enrollment

	Female	SES Index	Immigrant	Immigrant: first-generation	Immigrant: second-generation
	(1)	(2)	(3)	(4)	(5)
Early enrolled	0.045***	0.324***	-0.031***	-0.010***	-0.021***
	(0.004)	(0.013)	(0.002)	(0.001)	(0.001)
Late enrolled	-0.053***	-0.477***	0.567***	0.493***	0.074***
	(0.003)	(0.007)	(0.004)	(0.004)	(0.002)
No. Observations	1,255,947	1,255,947	1,255,947	1,255,947	1,255,947

Notes: each column shows the results of OLS regressions where the dependent variable are, respectively, a dummy variable indicating females, the SES Index, a dummy variable indicating immigrants, a dummy variable indicating immigrants of first and second generation (see definitions in Table 1). Early enrolled is a dummy variable indicating whether a student is enrolled in an earlier school cohort with respect to her age and the administrative cut-off rule; Late enrolled is a dummy variable indicating whether a student is enrolled in an older school cohort with respect to her age and the administrative cut-off rule. These two variables are contained in the SNV archives and derives from the schools' administrative archives. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.

Table 3
Baseline results

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Age Rank	-0.068***	-0.051***	-0.033***	-0.081*	-0.081*	-0.097**
	(0.003)	(0.003)	(0.004)	(0.048)	(0.049)	(0.041)
First stage F-stat.				781.42	815.11	1591.81
No. Observations	1,255,947	1,255,947	1,255,947	1,255,941	1,255,941	1,255,941
School by cohort FE	✓	✓	✓	✓	✓	✓
Individual level controls		✓	✓		✓	✓
Year of birth FE			✓			✓

Notes: the dependent variable is the probability of being bullied; all the regressions include as control variable students' age (defined in months). The age rank is calculated in percentile terms, as expressed in Equation 1. In the IV estimates (columns 4-6) the age rank and the age variables are instrumented using their corresponding theoretical counterparts, as obtained from Equation 4. Individual level control variables include gender, SES Index, immigrant status (first and second generation), class size and its square (for the definitions see Table 1); FE indicates fixed effects. The *First stage F-stat.* refers to the Kleibergen-Paap rk Wald F-statistics. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.

Table 4
Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age Rank	-0.271**	-0.094**	-0.108**	-0.107***	-0.097**	-0.093**	-0.118*
	(0.121)	(0.037)	(0.044)	(0.041)	(0.042)	(0.044)	(0.071)
First stage F-stat.	277.19	1742.50	1570.38	1570.34	818.68	456.68	597.99
No. Observations	606,369	1,255,941	1,243,636	1,243,636	1,255,941	1,255,941	1,255,946
All FE and controls	✓	✓	✓	✓	✓	✓	
Specifications:							
Discontinuity sample (Q1-Q4)	✓						
Age mean (year of birth by school)		✓					
Age variance (year of birth by school)			✓				
Age mean and variance				✓			
Quadratic term in age					✓		
Cubic term in age						✓	
School and cohort FE							✓

Notes: the table shows IV estimates in which the dependent variable is the probability of being bullied and the age rank and the age variables are instrumented using their corresponding theoretical counterparts; all the regressions (with the only exception of column 7) include the control variables and FE as specified for column 6 in Table 3. Specification in column 1 restricts the sample to students born in the first and fourth quarter of each calendar year; specifications in columns 2, 3 and 4, add as control variables, respectively, the mean, the variance and both the mean and the variance of the age variable, calculated at the year of birth by school level; specifications in columns 5 and 6 include, respectively, a quadratic and cubic term in age (instrumented with their quadratic and cubic theoretical counterparts); specification in column 7 include school and cohort fixed effects (two-way fixed effects model) instead of school-by-cohort fixed effects, and school-by-cohort time variable characteristics such as the share of females, first- and second-generation immigrants (the others FE and individual level control variables are left unchanged). The *First stage F-stat.* refers to the Kleibergen-Paap rk Wald F-statistics. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.

Table 5
Between-groups heterogeneous effects

	(1)	(2)	(3)	(4)
	Gender		*Immigrant status*	
	Males	*Females*	*Natives*	*Immigrants*
Panel A				
Age Rank	-0.081	-0.123**	-0.096**	-0.138
	(0.060)	(0.060)	(0.043)	(0.188)
First stage F-stat.	1354.98	1449.91	2049.45	136.54
Baseline probability	0.427	0.383	0.399	0.463
No. Observations	631,387	624,500	1,135,317	113,293
	Socio-economic status		*School time*	
	High	*Low*	*Half-day*	*Full-day*
Panel B				
Age Rank	-0.083	-0.129**	-0.057	-0.170**
	(0.061)	(0.061)	(0.049)	(0.078)
First stage F-stat.	1467.94	1064.31	1429.25	572.78
Baseline probability	0.391	0.420	0.391	0.435
No. Observations	618,181	636,505	838,759	416,641

Notes: the table shows IV estimates in which the dependent variable is the probability of being bullied and the age rank and the age variables are instrumented using their corresponding theoretical counterparts; all the regressions include the control variables and FE as specified for column 6 in Table 3. The groups of students with low or high socio-economic status are defined based on being, respectively, below or above the median value of the SES Index. See Table 1 for the definition of half- and full-day school time schedule. The *First stage F-stat.* refers to the Kleibergen-Paap rk Wald F-statistics. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.

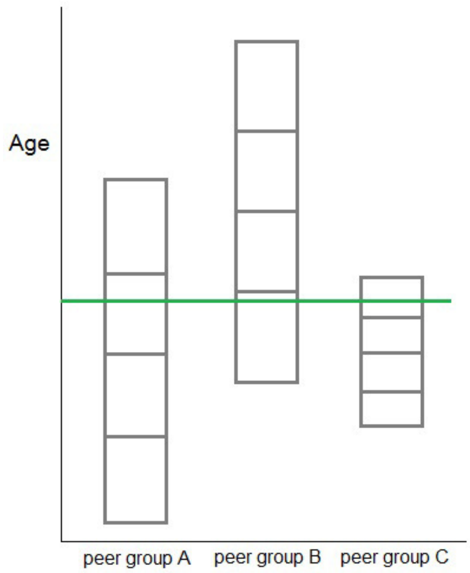
Table 6
Within-group heterogeneous effects

	(1)	(2)	(3)	(4)
	Gender		*Immigrant status*	
	Males	*Females*	*Natives*	*Immigrants*
Panel A				
Age Rank	-0.050	-0.042	-0.040	0.041
	(0.050)	(0.051)	(0.039)	(0.048)
First stage F-stat.	2630.84	1045.76	1332.55	443.11
Baseline probability	0.427	0.383	0.399	0.463
No. Observations	631,387	624,499	1,135,317	113,290
	Socio-economic status		*School time*	
	High	*Low*	*Half-day*	*Full-day*
Panel B				
Age Rank	-0.025	-0.032	-0.003	-0.036
	(0.030)	(0.028)	(0.027)	(0.027)
First stage F-stat.	1671.51	1884.60	1949.97	1852.46
Baseline probability	0.391	0.420	0.391	0.435
No. Observations	618,181	636,505	838,759	416,641

Notes: the table shows IV estimates in which the dependent variable is the probability of being bullied and the age rank (in this case, calculated within each group) and the age variables are instrumented using their corresponding theoretical counterparts; all the regressions include the control variables and FE as specified for column 6 in Table 3. The groups of students with low or high socio-economic status are defined based on being, respectively, below or above the median value of the SES Index. See Table 1 for the definition of half- and full-day school time schedule. The *First stage F-stat.* refers to the Kleibergen-Paap rk Wald F-statistics. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.

Appendix A Additional figures and tables

Figure A1
The ideal experiment: a graphical illustration



Notes: the three panels illustrate three peer groups which differ in their mean or variance of the age distribution (indicated on the horizontal axis). Peer group B displays the same age distribution as peer group A, except for the mean (vertical translation); peer group C displays the same age distribution as peer group A, except for the variance (compression). *Source*: based on Elsner and Isphording (2017a) and Murphy and Weinhardt (2014).

Table A1

School by cohort fixed effects estimates: extended results

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Age rank	-0.068***	-0.051***	-0.033***	-0.081*	-0.081*	-0.097**
	(0.003)	(0.003)	(0.004)	(0.048)	(0.049)	(0.041)
Age	0.003***	0.001***	-0.001***	0.002	0.002	0.004
	(0.000)	(0.000)	(0.000)	(0.004)	(0.004)	(0.003)
SES Index		-0.017***	-0.017***		-0.017***	-0.017***
		(0.000)	(0.000)		(0.000)	(0.000)
Female		-0.045***	-0.045***		-0.045***	-0.045***
		(0.001)	(0.001)		(0.001)	(0.001)
Immigrant (first-generation)		0.039***	0.037***		0.037***	0.035***
		(0.002)	(0.002)		(0.010)	(0.002)
Immigrant (second-generation)		0.043***	0.043***		0.043***	0.043***
		(0.002)	(0.002)		(0.002)	(0.002)
Class size		-0.002	-0.002		-0.002	-0.002
		(0.002)	(0.002)		(0.002)	(0.002)
Class size squared		0.000	-0.000		0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)
First stage F-stat.				781.42	815.11	1591.81
No .Observations	1,255,947	1,255,947	1,255,947	1,255,941	1,255,941	1,255,941
School by cohort FE	✓	✓	✓	✓	✓	✓
Individual level controls		✓	✓		✓	✓
Year of birth FE			✓			✓

Notes: the dependent variable is the probability of being bullied. The age rank is calculated in percentile terms, as expressed in Equation 1. In the IV estimates (columns 4-6) the age rank and the age variables are instrumented using their corresponding theoretical counterparts, as obtained from Equation 4. For the definitions of the control variables see Table 1; FE indicates fixed effects. The *First stage F-stat.* refers to the Kleibergen-Paap rk Wald F-statistics. Robust standard errors in parenthesis, clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. *Source*: based on Invalsi SNV.