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

Rosario Maria Ballatore† Marco Paccagnella‡ Marco Tonello§

This version: November, 2015

- Preliminary and incomplete -

Abstract

Using census data on four cohorts of Italian students enrolled in 5^{th} grade between 2010 and 2013, we investigate how age affects the probability of being bullied, distinguishing between the contribution of absolute and relative age. The absolute age effect is identified by means of an IV strategy that exploits the discontinuity in the probability of enrolling in a given school year generated by an end-of-year cut-off rule. In such a setting, theoretical school starting age is a valid instrument for absolute age. The relative age effect is identified exploiting cohort-to-cohort variation within schools in the average age of enrolled students, also instrumented by (average) theoretical school starting age. We find that only absolute age negatively affects the chances of being bullied, while relative age has no effect. This result implies that policies aimed at preventing bullying through changes in the age composition of schools and classes are unlikely to yield significant benefits.

JEL Classification: I21, J24, Z13

Keywords: bullying, school violence, relative age

*We thank Francesco Avvisati, Francesca Borgonovi, Lorenzo Cappellari, Marta De Philippis, Judith Pal, Claudio Lucifora, Michele Pellizzari, Paolo Sestito, and seminar participants at the LSE for useful comments and suggestions. We are indebted to Patrizia Falzetti and Valeria Tortora (INVALSI) for generous support in providing the data used in this paper. The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions they belong to. 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, Structural Economic Analysis Directorate, Law and Economics Division, Rome and CRELI-Catholic University Milan, E-mail: marco.tonello@bancaditalia.it.

1 Introduction

Violence and more general forms of misbehavior in schools are a widespread phenomenon, that is increasingly capturing the attention of policy-makers, parents, and schools' principals and administrators. Mühlenweg (2010) shows that the share of pupils reporting being victimized ranges between 30% (in Poland and Sweden) and 60% (in England and Spain). In the years between 2005 and 2013 in the US, the percentage of students who reported being bullied was on average 27 percent (Robers, Zhang, Morgan, and Musu-Gillette, 2015). The increasing attention to such issues is partly motivated by the dramatic increase in juvenile crime experienced by many industrialized countries since the late '90s, which is frequently linked to insufficient discipline and supervision by parents (outside schools) and teachers (in schools).

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. Being victimized has both short-term and longer-term consequences. Amongst the former, the literature has traditionally looked at outcomes such as anxiety, insecurity, low self-esteem, more frequent absences from school. The latter may include lower academic performance, and consequently worst labor market outcomes, as well as higher prevalence of anti-social behavior and all kind of psychological problems. Obviously, disentangling the direction of causality between such outcomes and victimization status is not an easy task. Recent attempts in this direction include Brown and Taylor (2008); Ponzo (2013), and Eriksen, Nielsen, and Simonsen (2014).

In this paper we investigate possible determinants of school victimization, following Leung and Ferris (2008); Mühlenweg (2010), and Ammermueller (2012). We focus in particular on the effect of age: our main contribution is to separately estimate the effect of *absolute* vs. *relative* age. Such focus is particularly relevant from a policy perspective. Contrary to other possible determinants of bullying, like psychological traits, cognitive ability, or the extent to which children live in troubled households, age is easily and immediately observable by school principals. Furthermore, age can be subject to manipulation by different actors. Parents can strategically choose whether to postpone or anticipate school enrollment. School principals have usually a say into such enrollment decisions, and, more importantly, can influence the process of class formation and school admission: by determining the age composition, they also determine pupils' relative age with respect to their classmates or schoolmates.

Using census data on four cohorts of Italian students enrolled in 5^{th} grade, we are able to separately identify the effect of absolute and relative age on the probability of being victim of bullying. The effect of relative age is estimated by exploiting cross-cohort, within-school variation in the average age of one's schoolmates. The effect of absolute age is identified with an IV strategy, where the instrument is the theoretical age at the time school starts. In Italy, enrollment in primary school usually occurs in the year a child turns 6. December 31^{st} thus generates a discontinuity in children's age at the time they start school (in September). Such discontinuity, however, is not sharp, because children born between January and April are allowed to enroll in the year they turn 5.

The rest of the paper is organized as follows. In section 2 we review the literature. Section 3 describes the institutional setting and the data used in the paper. Section 4 presents the empirical strategy, while Section 5 discusses the main results and some robustness checks. Section 6 extends the baseline analysis to other cognitive and non-cognitive outcomes. Section 7 concludes and derives policy implications.

2 Related literature

The literature on age effects in education has quite a long tradition, and it is way too large to be extensively covered here. We only notice that such literature has traditionally looked at the effects of age on cognitive outcomes, such as test scores (Bedard and Dhuey, 2006; Black, Devereux, and Salvanes, 2011; Cascio and Schanzenbach, forthcoming; Crawford, Dearden, and Meghir, 2010; Elder and Lubotsky, 2009). More recent papers have started to shift their focus to non-cognitive outcomes (Crafword, Dearden, and Greaves, 2015; Lubotsky and Kaestner, 2014), also as a channel to explain the effect of age on academic achievement (Pellizzari and Billari, 2012), and health (Dee and Sievertsen, 2015).

Despite its relevance in the process of human capital formation and the potentially very serious negative effects on subsequent outcomes, school bullying started to receive attention only recently, mainly thanks to both an increasing awareness of the seriousness of the problem, and to the availability of suitable micro-level datasets. However, convincing empirical evidence on the determinants of being a victim of bullying is still relatively scarce; credible estimates of the (causal) effects of being victimized are even less numerous.

Ammermueller (2012) finds that being a victim of bullying has a significantly negative impact on contemporary and later student performance (about 0.10 standard deviations in grade 4 and 0.18 in grade 8).[1] Eriksen, Nielsen, and Simonsen (2014) estimate the effect of being victim of bullying during the elementary school (i.e. at age 10-12) on the national assessments test scores taken at the end of the 9^{th} grade in Denmark finding that being victim of bullying decreases the test scores of about 1.14 standard deviations.

In the light of such evidence, identifying (and disentangling) the various potential determinants of victimization is indeed of crucial policy interest, as it would allow to design *ad hoc* interventions

[1]Similar findings are also shown by Ponzo (2013) for a sample of Italian students enrolled in the fourth (age 9) and in the eighth grade (age 13).

to counter and prevent their diffusion. Early studies in the psychological literature have traditionally focused on the analysis of individual level correlates of being victim of bullying, finding that personal traits such as low self-esteem, depression, loneliness, physical appearance are significantly associated to a higher probability of being bullied.[2] In a similar vein, Ammermueller (2012) finds that gender, social and migration background and the physical appearance of students are associated to being bullied, hurt or stolen from by peers. Leung and Ferris (2008) use data on self-reported violence acts of about 600 17-year-old males living in Montreal (Canada) in the 1990s to study how school victimization correlates with school size. The authors find that, after controlling for other plausible determinants of teenage violence, an increase of a hundred students enrolled in a school is associated with a one percent increase in the probability of suffering violence by a teenage peer. Their results seem also to point to a discontinuity in the effect of school size, such that only a very large school size matters.

Mühlenweg (2010) is probably the study most closely related to our paper. He exploits cross-country data from the Progress in International Reading Literacy Study (PIRLS) to estimate the effect of students' age on measures of self-reported school bullying and victimization. As students' age within grade is likely to be endogenous to the outcomes, especially when age rules are not strictly applied, official entry rule age based on the month of birth is used as an instrument for students' age within grade. The results show that the probability to suffer from victimization is reduced by 8 percentage points for older children, while younger children are significantly more likely to get something stolen, to be bullied and to be hurt. While this paper provides accurate evidence on the link between age and victimization at school, it does not allow for clear-cut policy conclusions. Indeed, the estimated age parameter incorporates both the effect of being older in an absolute sense and compared to their school peer. Without this distinction, no policy conclusions can be formulated. In our paper we exploit two independent sources of variation in the single student's age and in his/her schoolmates' age to disentangle these two channels at work.

3 Data and institutional setting

3.1 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 enroll in the first grade of the primary school the year they turn six. Up to 8^{th} grade, the school curriculum is identical for all students. Secondary education lasts three to five years, depending on the chosen track (vocational, technical, academic).

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

Starting from the 2009-10 school year, the Italian National Institute for the Evaluation of the Education System (Invalsi) carries out annual evaluations of students' attainment by means of the National Assessments Evaluation Surveys (SNV). The SNV includes a background questionnaire and a direct assessment of mathematics and the Italian language. It takes the form of an annual census, since it is compulsory for all schools and students attending second, fifth and sixth grade. For each grade, about 500,000 students participate to the assessment, which takes place in two different days (for the two subjects) in late May.

The SNV data contain therefore information on test scores, individual level and family background characteristics, that are collected from two main sources: (i) students' characteristics are directly compiled by the school administrative staff, drawing from school administrative records, and appear on each student's answer sheet; (ii) additional individual-level information on family, school and environmental characteristics are collected through a *Student Questionnaire* taken by 5th and 6th grade students the same day of one of the test (after finishing the exam). Parental background information are also available, such as mother's and father's place of birth (Italy, EU, European but non-EU, other non-European country), occupation and education level.

Crucially for our purposes, the *Student Questionnaire* for 5^{th} graders contains a full set of questions on students behavior at school and outside school. From these questions, we retrieve the information on students' victimization and other non-cognitive outcomes, such as self-concept, peer exclusion, and frequency of meetings with fiends outside the school.[3] Finally, we were able to gain access to the month of birth of each student, which we use to compute student's age in months.

3.2 Data and descriptive evidence

We exploit four waves of the SNV, corresponding to the school years from 2010/11 to 2013/14, and we focus on 5^{th} grade students for which we can retrieve information on victimization from the *Student Questionnaire*. The data include class and school identifiers, which allow 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 student's schoolmates. Moreover, using the universe of the students enrolled in a given grade entails a notable advantage for our identification strategy, as it makes it possible to precisely compute averages of variables at the school level (as all school peers are in the dataset, and not only a fraction of them, as in surveys sampling only students or classes) overcoming serious problems of attenuation bias (Micklewright, Schnepf, and Silva, 2012).

In line with existing works that exploit international surveys such as PIRLS or TIMMS (Mühlenweg, 2010; Ponzo, 2013), we identify students as victims of bullying whenever they answered positively to

[3] The variables describing non-cognitive outcomes have been constructed following the procedure suggested by Invalsi and described in Alivernini and Sestito (2014).

any of the following questions: (*i*) the student was beaten; (*ii*) the student was forced to do something against his or her will; (*iii*) the student was stolen things.[4] From the census of 5^{th} grade students we select the individuals born between 1998 and 2004, as those born before or after are more likely to be retained students or rare exceptions to the age admission rules. Moreover, we drop schools in which less than 10 students are enrolled in 5^{th} grade, as well as schools that are observed in one wave only. Finally we exclude all individuals who do not provide answer to the victimization question on the *Student Questionnaire*. Table 1 presents general descriptive statistics on the dependent and control variables using our final sample of about 1,300,000 students.

[Table 1 about here]

Almost 22 percent of the students declared having been victim of some sort of bullying at school. This figure is in line with existing evidence from other European countries: Ammermueller (2012) uses data form the Trend in International Mathematics and Science Study (TIMMS) for 4^{th} grade students in 2003 and finds that between 24 and 47 per cent of the students declared having been hurt or hit by a schoolmate in the month before the survey, while 12 to 32 percent reported having been stolen things. On average, students in our sample are about 10 years-old (121.6 months), while their theoretical starting age (calculated as of the beginning of primary school) is about 6.2 years (74.4 months). The table also reports the demographic characteristics (computed at the school level) that we use as control variables, i.e. the share of females, the share of natives, the share of first and second generation immigrants.[5]

4 Empirical strategy

The observed effect of age can be conceptually decomposed in four different parts. First, there is an effect from starting school earlier or later (i.e. the School Starting Age effect, SSA). Second, there is an effect from age at the time the test is taken (i.e. the Age At Test effect, AAT). Third, there is an effect from the total number of years of schooling (i.e. the Years Of Schooling effect, YOS). Lastly, there is a relative age effect (i.e. RAE), which comes from differences in individuals' age as compared to some reference peer group. The first three effects are linearly dependent: $AAT = SSA + YOS$, making it impossible to separately estimate the contribution of each single effect on some observed outcome (say, the probability of victimization). We label the cumulative effects of these three elements the *absolute age effect* (i.e. AAE). Relative age, while usually highly correlated with absolute age for obvious reasons, is not linearly dependent, which makes possible separate identification of its effect.

[4]Using alternative definitions, or defining three different outcome variables (one for each different question), do not change substantially our results.

[5]First generation immigrants are those born abroad, while second generation immigrants are born in Italy. Immigrant students are those whose both parents hold a non-Italian citizenship.

Our empirical strategy aims precisely at estimating the separate contribution of absolute and relative age on the individual probability of being victimized at school.

Following Elder and Lubotsky (2009), our starting point is the following model:

$$y_{isw} = \beta_0 + \beta_1 Age_{isw} + \beta_2 \overline{Age}_{sw(-i)} + \beta_3 X_{isw} + \beta_4 W_{sw} + \varepsilon_{isw} \tag{1}$$

in which y_{isw} is a dummy equal to one if student i in school s of wave w reported to be victim of bullying (zero otherwise), Age_{isw} is the student's age (in months), $\overline{Age}_{sw(-i)}$ is the average age of the student's schoolmates (in months), and X_{isw} and W_{sw} are individual and school level covariates. In order to highlight the absolute and the relative age component, equation 1 can be rewritten in the following way:

$$y_{isw} = \phi_0 + \phi_1 Age_{isw} + \phi_2 (Age_{isw} - \overline{Age}_{sw(-i)}) + \phi_3 X_{isw} + \phi_4 W_{sw} + \eta_{isw} \tag{2}$$

such that $\phi_1 = \beta_1 + \beta_2$ captures the absolute age effect (AAE), and $\phi_2 = -\beta_2$ captures the relative age effect (RAE).

Threats to the identification of the parameters of interest (ϕ_1 and ϕ_2) come from two main sources.

The first threat to identification comes from the endogenous choice of parents about the timing of birth and of school enrollment. Children are required to start primary school in the year in which they turn 6, so that December 31^{st} generates a natural cut-off date that changes in a discontinuous way both the relative position of a student within the school and the actual age at the time of the test. The rule is to some extent flexible, as it allows children born between January and April to enroll earlier, i.e. in the year they turn 5 (this is what we label *early enrollment*).[6] Students are also allowed to postpone enrollment (so-called *late enrollment*), but this latter case is much less frequent (as illustrated in Figure 1), and is usually motivated by specific problems such as a disability or difficulties with the instruction language for non-native students.

[Figure 1 about here]

The enrollment rule generates the theoretical starting age function expressed by Equation 3, which is a discontinuous function in the month of birth:

$$f_{isw} = 72 - (9 - m_{isw}) \tag{3}$$

[Figure 2 about here]

[6] The administrative rules are detailed in the Ministry of Education Regulations. For the academic year 2009-10 see the *Circolare Ministeriale No. 4/2009.*

Figure 2 displays the relationship between the theoretical starting age function (as expressed by Equation 3) and the school starting age observed in the data: compliance with the theoretical enrollment rule based on the 31st December cut-off is good, although not perfect, especially for the students 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 that could observe several characteristics of the child that are unknown to the researcher and that are plausibly correlated with the outcomes of interest. The fact that parents can endogenously select their children into early or late enrollment is also illustrated in Table 2, which shows that, for instance, female and children with a higher index socio-economic status (henceforth labeled *ESCS*) are more likely to enroll earlier, while immigrants are generally more likely to postpone enrollment.

[Table 2 about here]

To tackle the potential endogeneity of Age_{isw}, we exploit the 31^{st} December cut-off as a source of exogenous variation, and, given that compliance is not perfect, we instrument it with its theoretical counterpart (i.e. f_{isw}). 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. In the baseline analysis we exploit the full sample of students born in all months of the year, while in the robustness checks we will consider a specification that focuses on what we define as the discontinuity sample, constituted by all students born close to the cut-off (i.e. in December and January).

The second potential source of endogeneity is the fact that students are not randomly assigned to a pool of schoolmates. This is likely to happen for two main reasons: because of age-based sorting of students in different schools and because of manipulation in the process of class formation by school administrators. We abstract from the latter issue by focusing on the school level, although we will move to the class level as a robustness check. The former concern is addressed by controlling for school fixed effects, under the assumption that, within school, cohort to cohort variation in the average age of the student body is as good as random. However, cohort to cohort variation in the average age of peers is necessary, but not sufficient, for separate identification of the relative age effect (i.e. ϕ_2) (Cascio and Schanzenbach, forthcoming). Because of the endogenous choices of the parents to enroll children late or early with respect to the natural cut-off, the average of peers' age also embeds this selection process and this potential source of endogeneity. For this reason, we instrument $\overline{Age}_{sw(-i)}$ with $\overline{f}_{sw}$, i.e. the school level average of f_{isw}.

5 Results

5.1 Baseline results

We estimate linear probability models of several variants of Equation 1, with robust standard errors clustered at the school level.[7]

[Table 3 about here]

Table 3 presents the results from a simple OLS estimation. Columns 1 and 2 show the results for what we define to as the *conventional approach*, i.e. the approach usually taken in the previous literature that focuses on either absolute or relative age effects. The results show that both absolute and relative age are negatively correlated with the probability of being bullied at school. In columns 3 to 5 we jointly estimate both effects: column 3 reports the raw correlations, while column 4 and 5 add, respectively, the school level control variables (as listed in Table 1) and school fixed effects. Panel A presents the coefficient estimates of the OLS regressions, while in Panel B we show the Absolute (AAE) and the Relative Age Effect (RAE), given by, respectively, $\phi_1 = \beta_1 + \beta_2$ and $\phi_2 = -\beta_2$ in Equation 2. Focusing on column 5, our baseline specification which includes both the school characteristics and the school FE, the OLS estimates uncover a negative and statistically significant coefficient associated to absolute age and no significant effect on relative age. However, the endogeneity problems illustrated so far do not allow us to derive any conclusion on the direction of the causal link between absolute age, relative age and being bullied.

[Table 5 about here]

At least from a qualitative point of view, the results of the OLS regressions are confirmed when we apply the IV strategy illustrated in the previous section. The conventional estimates in columns 1 and 2 of Table 5 still present a negative and statistically significant effect larger, in absolute terms, with respect to the corresponding conventional OLS estimates (Table 3, columns 1 and 2). This is likely to occur because the OLS estimates are biased upward as parents anticipate the enrollment of children if they think they are ready for schools (and thus less likely to be victimized even if they are the youngest), or *viceversa* because parentes postpone the enrollment of pupils not ready to attend school (which are likely to be victims of bullism despite they are the oldest).

The negative effect of absolute age and the absence of effects of relative age is confirmed across the different specifications, and the estimated coefficient for absolute age is three times larger than the corresponding OLS estimate. The estimates for the baseline IV specification including the control

[7] 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 if we adopt either a logit or a probit specification.

variables and the school fixed effects (column 5) suggest that, while relative age does not have any statistically significant effect, each additional month of (absolute) age reduces the probability of being victimized by 0.3 percentage points. Such an effect is not negligible: if we cumulate the estimated effects over the age difference in terms of months between a child born in January and one born in October, we obtain that the child born in January faces a probability of being bullied which is 3 percentage points lower than the child born in October.

The negative effect on own (absolute) age mirrors previous results in the literature and it can be easily rationalized with older children having developed better non-cognitive (and cognitive) skills, so that being older implies *per se* a lower probability of suffering bullying at school. The lack of effect for the relative age is new in the literature and somehow less intuitive: it implies that being younger than schoolmates does not affect *per se* the probability of being victim of bullying. Our results are not in contrast to those in Mühlenweg (2010), and represent a step forward in the understanding of the age effects on students victimization. Mühlenweg (2010) finds that the probability of victimization is reduced for older students. In turn, her results are similar to our findings for the conventional estimates of relative age effects (column 2).

[Table 4 about here]

Finally, table 4 shows the results of the first-stage regressions for each endogenous variable (Age_{isw} in Panel A and $\overline{Age}_{sw(-i)}$ in Panel B). Column 3 shows that the theoretical age (f_{isw}), calculated following the rule expressed in Equation 3, is a strong predictor of both absolute and relative age; the school level average of the theoretical age ($\overline{f}_{sw}$), is again strongly correlated with both absolute and relative age. The first-stage F-statistics show that there are no worries about the strength of the instruments according to thresholds defined by Stock and Yogo (2005).

5.2 Heterogeneous effects

We further investigate whether our results show important heterogeneities with respect to observable characteristics of the students, such as gender, immigrant status and family background (as proxied by the ESCS indicator). Table 6 shows that our main results are remarkably consistent across different groups: we find a negative and statistically significant AAE for all the groups, while the RAE is never statistically different from zero. The effect of absolute age, however, tends to be much stronger for non-natives and for males.

[Table 6 about here]

5.3 Robustness checks

Before extending our analysis to achievement and non-cognitive skills, we present some robustness checks for our main results. A first potential threat to our identification strategy comes from the fact that we cannot flexibly control for quarter of birth in the baseline specification given that it is highly correlated with our instrumental variables. An extensive literature has shown that the quarter of birth correlates with both cognitive and non-cognitive skills, early and later in life. To tackle this issue, we restrict the sample to children born in December or January (i.e. what we defined to as the *discontinuity sample*): for these children there should be no variation in unobservable characteristics related to the quarter of birth. The results are presented in Table 7 and do not change significantly from the baseline 2SLS specification (i.e. Table 5, column 5). We only find a slightly larger effect of absolute age in the discontinuity sample, which is what one would expect given the focus on those students that experienced the sharpest discontinuity in the age rule determined by the enrollment cut-off date.

[Table 7 about here]

A second threat to our identification strategy, potentially inducing an underestimation of the RAE, is the choice of schoolmates rather than classmates as the reference peer group. The choice of looking at schoolmates instead of classmates is motivated by two main reasons. First, descriptive evidence from school victimization studies shows that most of the bullying cases actually occur in the school premises (such as the gym, the school cafeteria or the canteen) but outside the classroom. The 2015 report on victimization at school in the US (Robers, Zhang, Morgan, and Musu-Gillette, 2015) shows that in 66 percent of the cases episodes of bullying are declared to occur outside the classroom. Second, the allocation of students across classes within schools is endogenous and generally determined by the choices of school principal and administrative staff, which could very well be based on students characteristics such as age and citizenship, for which it is impossible to convincingly control for in absence of a randomized experiment.[8]

However, at least in principle, students interactions are stronger and more frequent within the classroom and among classmates, and this is probably especially true for younger students enrolled in primary school. To partly address the problem of non-random class formation, we present estimates in which, while computing relative age with respect to classmates rather than schoolmates, we restrict the sample to schools for which we can not reject the hypothesis of random allocation of students in classes (based on students' observable characteristics) (Contini, 2013). More precisely, we only keep those schools for which a Pearson's chi-squared test of random allocation (according to age) of students across classes

[8] For example, Cascio and Schanzenbach (forthcoming) overcome this problem exploiting the random allocation scheme of the STAR Project.

within a given school cannot be rejected (Ammermueller and Pischke, 2009). The results presented in Table 7, columns 2 and 3, are remarkably consistent with our baseline estimates.

6 Extension: cognitive and non-cognitive outcomes

Previous results in the literature on the age effects on educational outcomes have prevalently focused on cognitive outcomes, such as test scores, generally finding positive cognitive effects from the interaction with older peers, which are usually interpreted as counterbalancing the negative effects given from absolute age (Bedard and Dhuey, 2006; Black, Devereux, and Salvanes, 2011; Cascio and Schanzenbach, forthcoming; Crawford, Dearden, and Meghir, 2010; Elder and Lubotsky, 2009). Less consensus is found concerning non-cognitive outcomes, as the behavioral channels underlying the reduced form effects can be hardly identified (Pellizzari and Billari, 2012).

In this section we exploit the richness of our dataset both in terms of cognitive measures (i.e. test scores) and non-cognitive measures that can be retrieved from the *Student Questionnaire* to provide evidence on the effects of absolute and relative age on a wider array of outcomes. The aim of this exercise is twofold. On the one hand, understanding whether relative age might affect achievement and non-cognitive outcomes provides suggestive evidence concerning the potential cognitive and behavioral mechanisms underlying the reduced form results that we find for the victimization outcome. On the other hand, the analysis can benchmark our results with the existing literature on absolute and relative age effects, thus providing additional pieces of evidence to shade light on the effects of relative age on non-cognitive outcomes.

[Table 8 about here]

Table 8 contains the estimates for the effects of relative and absolute age on achievement and non-cognitive outcomes. As for our previous tables, Panel A shows the estimates from the regressions, while in Panel B we adjust them according to the AAE and RAE framework (Equation 2). Achievement is measured by the test score taken for literacy skills,[9] while we focus on three main non-cognitive outcomes that are plausibly related to victimization: feeling excluded by peers, frequency of meeting friends outside the school, self-concept (Alivernini and Sestito, 2014).

The results for achievement (column 1) uncover a positive effect of absolute age and a negative effect of relative age. The RAE is also much smaller as compared to the AEE but still strongly statistically significant: increasing absolute age by one month would determine an increase in the achievement of about 0.03 standard deviations, while an increase by one month in the relative age would decrease achievement by 0.005 standard deviations. This result is in line with previous findings in the few studies

[9]The results (available upon request) do not change if we consider the numeracy test.

looking at relative age effects (Cascio and Schanzenbach, forthcoming). Moreover, the fact that we do find a relative age effect for the achievement outcome is also also reassuring about the ability of out identification strategy to correctly capture the RAE.

The benefits of the exposure to older peers, however, could work through other channels that improve non-cognitive, rather than cognitive, development. The results for the non-cognitive outcomes are mixed: relative age does not appear to have statistically significant effects on any the three non-cognitive outcomes, while absolute age positively affects self-concept and negatively affects the feeling of exclusion by peers. Even in these cases the signs are those expected: self-concept increases the older is the student, while the feeling of being excluded by peers decreases. These results seem to indicate that behavioral outcomes and victimization are not affected by relative age *per se*, while it is the absolute age effect that matters. This is a crucial indication for policy making.

7 Concluding remarks

Being bullied at school is a growing phenomenon that has only recently grasped the attention of the policy makers. Being bullied at school has been found to have both short and long run negative effects on cognitive and non-cognitive skills, but also on labor market performance and health (Bowes, Joinson, Wolke, and Lewis, 2015). Recent studies make considerable contribution to this growing body of the literature adopting more sophisticated identification strategies (Ammermueller, 2012; Eriksen, Nielsen, and Simonsen, 2014; Ponzo, 2013), and confirming the negative effects in the short term 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 not clear from these works which is the direction of the causality link between individual traits and being bullied.

Nevertheless, understanding the determinants of school victimization is of crucial importance for the design of policies aimed at countering bullying at school. This paper aims to fill this gap in the literature by estimating the effects that one potential important channel might have on the probability of being bullied at school, notably relative age. Disentangling the separate contribution of the relative age affect, as compared to the absolute age effect, is of crucial importance to determine whether being bullied at school is significantly affected by the exposure to older peers. We defined relative age as the average difference between each student age (in months) and the age of his/her school and grade mates. To this purpose, we exploit a rich dataset covering the census of 5^{th} grade students enrolled in Italian schools in the school years from 2010/11 to 2013/14 (about 1,300,000 children), containing self-reported information about the bullying episodes suffered by each respondent.

We find that each additional month of age reduces the probability of being victimized by 0.3 percentage points, while relative age does not have any statistical significant effect. The absolute age effect is not negligible. For instance, if we cumulate the estimated effects over the age difference in terms of months between a child born in January and one born in October, we obtain that the child born in January faces a probability of being bullied which is 3 percentage points lower than the child born in October. We also find that the relative age effect does not play any role in any of the non-cognitive outcomes related to victimization (i.e. peer exclusion, self-concept and frequency of meeting friends) suggesting that relative age does not affect victimization through these alternative behavioral channels.

Our results suggest that policies aimed at changing the age composition of classes or schools cannot reduce victimization. Such policies would not have any effect on non-cognitive outcomes, such as self-concept or peer exclusion, through the relative age effect, while only affecting students' achievement.

References

ALIVERNINI, F., AND P. SESTITO (2014): "Motivazioni e percezioni degli studenti Italiani," Working Paper 21, Invalsi.

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

AMMERMUELLER, A., AND J.-S. PISCHKE (2009): "Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study," *Journal of Labor Economics*, 27(3), 315–348.

BEDARD, K., AND E. DHUEY (2006): "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects," *The Quarterly Journal of Economics*, 121(4), 1437–1472.

BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2011): "Too Young to Leave the Nest? The Effects of School Starting Age," *The Review of Economics and Statistics*, 93(2), 455–467.

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," *British Medical Journal*, 350, h2469.

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.

CASCIO, E., AND D. W. SCHANZENBACH (forthcoming): "First in the Class? Age and the Education Production Function," *Education Finance and Policy*.

CONTINI, D. (2013): "Immigrant background peer effects in Italian schools," *Social Science Research*, 42(4), 1122 – 1142.

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.

CRAWFORD, C., L. DEARDEN, AND C. MEGHIR (2010): "When you are born matters: the impact of date of birth on educational outcomes in England," IFS Working Papers W10/06, Institute for Fiscal Studies.

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.

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.

ERIKSEN, T. L. M., H. S. NIELSEN, AND M. SIMONSEN (2014): “Bullying in Elementary School,” *The Journal of Human Resources*, 49(4), 839–871.

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.

LEUNG, A., AND S. J. FERRIS (2008): “School Size and Youth Violence,” *Journal of Economic Behavior and Organization*, 65, 318–333.

LUBOTSKY, D., AND R. KAESTNER (2014): “Effects of Age at School Entry on Child Cognitive and Behavioral Development,” Discussion paper, University of Illinois.

MÜHLENWEG, A. M. (2010): “Young and Innocent: International Evidence on Age Effects within Grades on Victimization in Elementary School,” *Economics Letters*, 109(3), 157–160.

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.

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,” NCES 2015-072/NCJ 248036, 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.

STOCK, J. H., AND M. YOGO (2005): “Testing for Weak Instruments in Linear IV Regression,” in *Identification and Inference for Econometric Models*, ed. by D. Andrews, pp. 80–108. Cambridge University Press, New York.

Figures

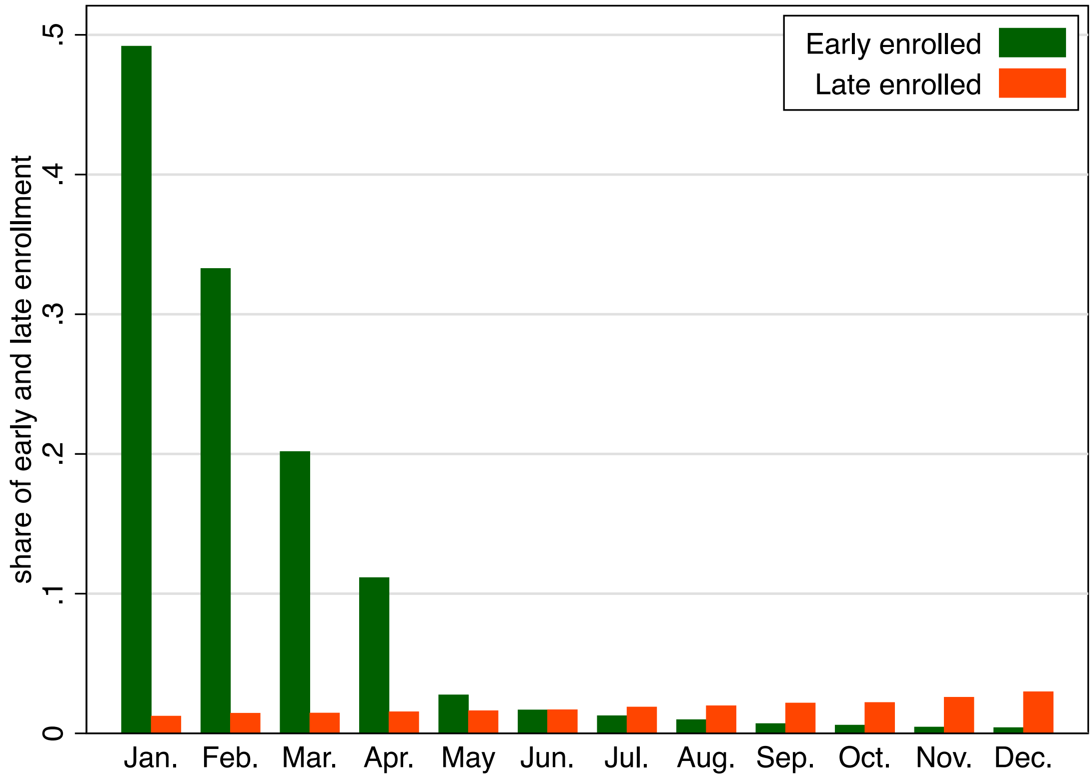

Figure 1: Early and late enrollment

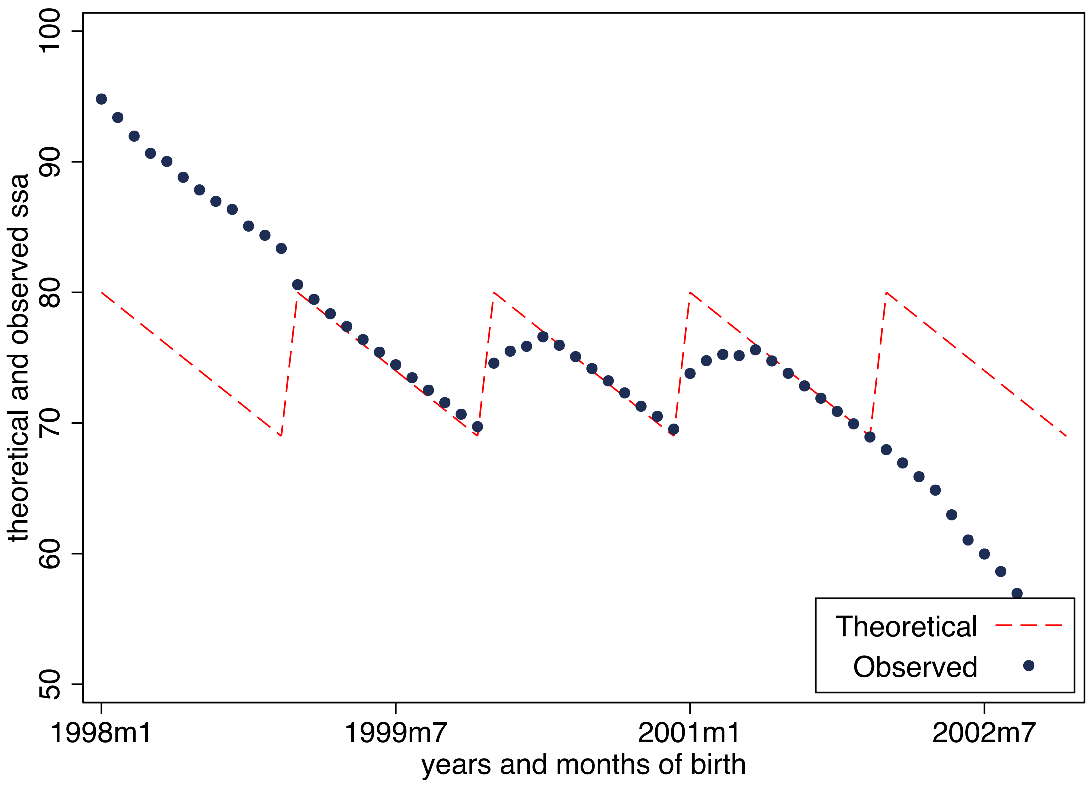

Figure 2: Theoretical and observed school starting age

Tables

Table 1: Descriptive statistics

	mean	sd
Dependent variable:		
Bullied (dummy)	0.22	0.41
Control variables:		
Age (in months)	121.65	4.62
Theoretical starting age (in months)	74.44	3.42
Escs	0.13	1.02
Share of female	0.50	0.50
Share of natives	0.91	0.29
Share of 1st gen. imm.	0.04	0.21
Share of 2nd gen. imm.	0.05	0.22
Additional outcome variables:		
Test score (language)	77.49	13.53
Freq. friends	3.27	0.90
Social exclusion (dummy)	0.26	0.44
Self concept	0.00	1.37
# of schools	14,077	
# of classes	27,190	
# of students	1,292,293	

Notes: Escs refers to the socio-economic indicator of family background. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 2: Early and late enrollment

	Escs (1)	1(female) (2)	1(immigrant) (3)
1(Early enrolled)	0.1623***	0.0537***	-0.0149***
	(0.003)	(0.001)	(0.001)
1(Late enrolled)	-0.4498***	-0.0525***	0.5073***
	(0.005)	(0.002)	(0.003)
Observations	1,292,293	1,292,293	1,292,293

Notes: Escs refers to the socio-economic indicator of family background. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 3: OLS results

	(1)	(2)	(3)	(4)	(5)
Panel A					
Age_{isw} ($\hat{\beta}_1$):	-0.0016***		0.0000	-0.0012***	-0.0016***
	(0.000)		(0.000)	(0.000)	(0.000)
$\overline{Age}_{sw(-i)}$ ($\hat{\beta}_2$):			0.0063***	0.0044***	0.0005
			(0.001)	(0.001)	(0.000)
$(Age_{isw} - \overline{Age}_{sw(-i)})$:		-0.0015***			
		(0.000)			
Panel B					
AAE ($\hat{\beta}_1 + \hat{\beta}_2$):			0.0063***	0.0033***	-0.0010**
			(0.001)	(0.001)	(0.000)
RAE ($-\hat{\beta}_2$):			-0.0063***	-0.0044***	-0.0005
			(0.001)	(0.001)	(0.001)
Controls	y	y	n	y	y
School fe	y	y	n	n	y
Observations	1,292,293	1,292,293	1,292,293	1,292,293	1,292,293

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 4: First stage

	(1)	(2)	(3)
Panel A	First stage on Age_{isw}		
f_{isw}	0.4648***	0.4680***	0.4676***
	(0.003)	(0.004)	(0.004)
$\overline{f}_{sw(-i)}$	0.0122*	-0.0079	-0.0380***
	(0.007)	(0.006)	(0.007)
F stat:	8962	8941	8763
Controls	n	y	y
School fe	n	n	y
Observations	1,292,293	1,292,293	1,292,293
Panel B	First stage on $\overline{Age}_{isw}$		
f_{isw}	-0.0026***	-0.0029***	0.0070***
	(0.000)	(0.000)	(0.000)
$\overline{f}_{sw(-i)}$	1.2356***	1.1896***	1.1888***
	(0.038)	(0.042)	(0.043)
F stat:	417.7	423.4	354.2
Controls	n	y	y
School fe	n	n	y
Observations	1,292,293	1,292,293	1,292,293

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 5: IV results

	(1)	(2)	(3)	(4)	(5)
Panel A					
Age_{isw} ($\hat{\beta}_1$):	-0.0029***		-0.0036***	-0.0032***	-0.0029***
	(0.000)		(0.000)	(0.000)	(0.000)
$\overline{Age}_{sw(-i)}$ ($\hat{\beta}_2$):			0.0003	0.0001	-0.0003
			(0.000)	(0.000)	(0.000)
$(Age_{isw} - \overline{Age}_{sw(-i)})$:		-0.0025***			
		(0.000)			
Panel B					
AAE ($\hat{\beta}_1 + \hat{\beta}_2$):			-0.0033***	-0.0032***	-0.0032***
			(0.001)	(0.001)	(0.001)
RAE ($-\hat{\beta}_2$):			-0.0003	-0.0001	0.0000
			(0.000)	(0.000)	(0.001)
Controls	y	y	n	y	y
School fe	y	y	n	n	y
Observations	1,292,293	1,292,293	1,292,293	1,292,293	1,292,293

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 6: Heterogeneity of results

	Gender		Imm. status		Fam. background	
	male	female	natives	imm.	high	low
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Age_{isw}	-0.0040***	-0.0017***	-0.0025***	-0.0060***	-0.0028***	-0.0031***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
$\overline{Age}_{sw(-i)}$	-0.0000	-0.0007	-0.0002	-0.0013	-0.0004	-0.0004
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Panel B						
AAE:	-0.0041***	-0.0024***	-0.0028***	-0.0074***	-0.0033***	-0.0035***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
RAE:	0.0000	0.0007	0.0002	0.0013	0.0005	0.0004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	y	y	y	y	y	y
School fe	y	y	y	y	y	y
Obs	649,805	642,479	1,169,132	121,374	669,404	622,750

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 7: Robustness checks

	Discontinuity sample	Classmates Peers	
		whole sample	"random" sample
	(1)	(2)	(3)
Panel A			
Age_{isw} ($\hat{\beta}_1$):	-0.0030***	-0.0027***	-0.0029***
	(0.001)	(0.000)	(0.000)
$\overline{Age}_{sw(-i)}$ ($\hat{\beta}_2$):	-0.0012		
	(0.001)		
($\overline{Age}_{cw(-i)}$):		-0.0005	-0.0006
		(0.001)	(0.001)
Panel B			
AAE ($\hat{\beta}_1 + \hat{\beta}_2$):	-0.0042***	-0.0032***	-0.0035***
	(0.001)	(0.001)	(0.001)
RAE ($-\hat{\beta}_2$):	0.0012	0.0005	0.0006
	(0.001)	(0.001)	(0.001)
Controls	y	y	y
School fe	y	y	y
Observations	212,025	1,292,293	1,111,518

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.

Table 8: Other outcomes

	Test scores (1)	Peer excl. (2)	Freq. friends (3)	Self concept (4)
Panel A				
Age_{isw} ($\hat{\beta}_1$):	0.0259***	-0.0021***	-0.0003	0.0157***
	(0.000)	(0.000)	(0.001)	(0.001)
$\overline{Age}_{sw(-i)}$ ($\hat{\beta}_2$):	0.0050***	0.0007	0.0028	0.0013
	(0.002)	(0.001)	(0.002)	(0.002)
Panel B				
AAE ($\hat{\beta}_1 + \hat{\beta}_2$):	0.0309***	-0.0014***	0.002	0.0170***
	(0.003)	(0.001)	(0.003)	(0.002)
RAE ($-\hat{\beta}_2$):	-0.0050***	-0.0007	-0.0028	-0.0013
	(0.002)	(0.001)	(0.003)	(0.002)
Controls	y	y	y	y
School fe	y	y	y	y
Observations	1,292,293	1,292,293	1,292,293	1,292,293

Notes: robust standard errors clustered at the school level. Asterisks denote statistical significance at the * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ levels. **Source**: based on SNV Invalsi for the school years 2010/11, 2011/12, 2012/13, 2013/14.