The older the bolder: Does relative age among peers influence children's confidence and risk attitudes?

Lionel Page*, Dipanwita Sarkar, and Juliana Silva Goncalves

Queensland University of Technology

March 27, 2015

Abstract

Relative age at school has been found to have long lasting consequences on a wide range of outcomes. We investigate whether the relative age position of young people in their cohort of peers plays a role in shaping their behavioural traits. We ran a laboratory experiment in the field over two states in Australia to elicit the preferences for competition, overconfidence, risk and ambiguity aversion, and professional aspiration levels of over 500 high school students who are the very oldest or very youngest in their class. Across all the behavioural variables investigated we find only mild evidence of an effect of relative age on these psychological traits.

1 Introduction

A large body of evidence points to the fact that children who are the oldest in their cohort at school tend to have higher educational achievement than the students who are the youngest (Datar, 2006; Puhani and Weber, 2007; McEwan and Shapiro, 2008; Smith, 2009; Elder and Lubotsky, 2009; Grenet, 2009; Crawford et al., 2014). This relative age effect (RAE) is also observed for non-cognitive skills (Thompson et al., 2004; Dhuey and Lipscomb, 2008) and for sporting performance (Musch and Grondin,

*We thank Martin Kocher and Guido Friebel for helpful comments. We also thank participants at the workshop on self-control, self-regulation and education in Aarhus, at the seminar series of the TrygFonden's Centre for Child Research (Aarhus University) and at the behavioural economics brown bag seminar series of the University of Munich (LMU) for interesting discussions. This research benefited from funding from the Australian Research Council (DE120101270) and the Queensland Department of Employment, Economic Development and Innovation Queensland (Smart Futures Fund).

2001). The existing evidence suggest that this effect can have long-lasting consequences, influencing the choice of professional careers and labour market outcomes (Bedard and Dhuey, 2006; Fredriksson and Öckert, 2006; Bedard and Dhuey, 2007; Grenet, 2009; Du et al., 2012; Crawford et al., 2014; Müller and Page, 2015). There is no clear consensus in the literature on which mechanisms drive the observed RAE. The debate is centred on whether the RAE is mainly explained by a maturity advantage of being older in a cohort, or whether the advantage stems from the *relative position* in the age distribution of the peer group (Musch and Grondin, 2001; Crawford et al., 2014).

Behavioural research has pointed to the role of relative position of young people in their peer group. Manski (2000), suggests that agents may learn from their peers about what they can achieve ("observational learning"). Battaglini et al. (2005) have shown that individuals' self-confidence may benefit from being paired with peers slightly weaker than them. Evidence put forward by Murphy and Weinhardt (2014) suggests that students' ordinal rank in primary school has lasting consequences on their success in high school. The effect of rank on self-confidence of children at an early age appears to be the main contending explanation for this effect and Murphy and Weinhardt (2014) suggest that the same mechanism may be at play in the relative age effect. In line with this suggestion, Crawford et al. (2014) find an effect of relative age on students' views about their scholastic competence at grade 8.

Past research has also found that children who are relatively old tend to be more confident and have higher self-esteem (Fenzel, 1992; Thompson et al., 1999, 2004). They are also less likely to suffer from school victimisation (Mühlenweg, 2010), and less likely to suffer from psychological disorders (Goodman et al., 2003; Morrow et al., 2012). In addition, relatively old students have been found to be more likely to be involved in leadership activities (Dhuey and Lipscomb, 2008). In a recent study looking at the US Congress, Müller and Page (2015) find that individuals who were relatively old at school are more likely to be among the leading politicians in the US. The authors conjecture that having been relatively old at school may have helped them build self-confidence and leadership skills. Taken together, these findings suggest that psychological mechanisms may constitute an important factor in explaining the RAE, with the possibility that younger children form lower expectations and confidence from interacting with older peers.

In the present study, we investigate the possible effect of young children's relative age among their peers at school on a wide range of preferences and attitudes which are relevant for economic behaviour. Specifically, we ran a laboratory experiment in the field[1] to elicit preferences for competitive environments, risk and ambiguity attitudes,

[1] Also called "artefactual field experiment" by Harrison and List (2004).

and professional aspirations of 661 high school students enrolled in grades 8 and 9 (aged 13 to 14 years old). We recruited students who were born within a two month window of the cut-off date defined to enter school.[2] Our experiment was conducted in 38 Australian schools across two states with different cut-off dates. We chose students in the age range 13 to 14 years old since recent literature suggest preferences of individuals in this age range to be stable (Heckman, 2007; Sutter et al., 2013; Harbaugh et al., 2002). Overall we find that the effect of relative age on behavioural variables are at best small.

This paper contributes to two strands of literature. First, it contributes to the literature on the RAE by investigating pathways which could drive this effect. Our study goes at the heart of the debate by studying whether younger and older children in a cohort differ on a wide range of behavioural traits such as self-confidence, risk attitude, preference for competition and professional aspirations. We find only limited evidence of an effect of relative age on such traits. From our results, we argue that differences in the behavioural traits we measure are unlikely to be important in explaining the long lasting differences in achievements between the relatively young and relative old in a cohort of students.

Second, and more broadly, this paper contributes to the growing literature in education looking at the effect of peers on students' behavioural outcomes and non-cognitive skills (Koch et al., 2014). To the best of our knowledge, this is the first study that uses experimental data in order to identify the importance of relative age of young children among their peers in moulding their behavioural traits. Our results indicate that if such effects exist, they may run through more subtle channels than the range of economic preferences and behaviour that are usually considered in the literature.

The remainder of this paper is structured as follows: Section 2 presents the empirical strategy, Section 3 describes our general experimental setup and participant pool, Section 4 presents our results and Section 5 discuss the results and concludes the study.

2 Empirical strategy

Our identification strategy builds on the regression over discontinuity methodology from Lee and Lemieux (2010). Such an approach has previously been used in the study of the relative age effect on educational achievement (Smith, 2009; Crawford et al., 2014). Here we use this method to investigate whether we observe relative age differences in economic preferences, with a sample of high school students born less than 61 days apart from each other and placed in two different school cohorts (grades 8 and 9) depending on the position of their birthday relative to the school entry cut-off date. We use this

[2] In most countries, children are eligible to start compulsory schooling conditional on reaching a certain age, by a defined cut-off date. This cut-off date can be defined nationally, or at the state level.

sample as the setting for a quasi experiment to compare the preference of the students relatively old in grade 8 and of the students relatively young in 9.

Specifically, we estimate the model:

$$y_{ij} = \beta_0 + \beta_1 ra_i + \beta_2 state_i + \mathbf{X}_i\beta_3 + \mathbf{S}_{ij}\beta_4 + \epsilon_{ij} \tag{1}$$

where y_i represents the outcome of interest for participant i, ra is a binary variable equal to 1 if the student is relatively old within the classroom and 0 if he or she is relatively young, and $state$ is an indicator variable for the state where the student is attending school (Queensland or New South Wales). The vectors $\mathbf{X}$ and $\mathbf{S}$ represent individual and school level control variables respectively. Individual level control variables include gender and academic scores. School level covariates control for whether the school is public or private, and located in a regional or metropolitan area.

As pointed out by Crawford et al. (2014), such an estimation cannot identify exactly the relative age effect, the paramater of interest is potentially also incorporating the effect of being in grade 9 rather than grade 8. Calling β_1^{qe} the coefficient estimate from this quasi-experimental setting, we have: $\beta_1^{qe} = f(relative\ age, grade)$.[3] Previous research has suggested that preferences are stable for adolescents by the age of 10 years old (Sutter et al., 2013; Harbaugh et al., 2002),[4]. Therefore it is plausible to assume that preferences will be stable when moving from one grade to the other. We would therefore expect the grade effect to be negligible and therefore our quasi-experimental estimate to mainly recover the relative age effect.[5] Calling β^{ra} the causal effect of relative age effect, we therefore expect that $\beta_1^{qe} = f(relative\ age, grade) \approx \beta^{ra}$.

We therefore run our first estimations on the subsample of students who are either the youngest in grade 9 or the oldest in grade 8. This sample contains students who were born very close in time but ended up having very different trajectories due to the position of their birthdate relative to their state cut-off dates. In Queensland these students were born either in December 1999 (grade 9) or in January 2000 (grade 8). In New South Wales, they were either born in July 2000 (grade 9) or August 2000 (grade 8). There is a well known issue of a potential sample selection bias due to imperfect compliance with school entry rules. Typically, students born before the cut-off dates are sometimes delayed in their school entry (so called "red shirting"). In our sample this means that some of the students born in December 1999 in Queensland will actually be in grade 8, as

[3] An effect of being in a given grade can itself come from an effect of the age at starting school and the effect of an additional year of schooling (Crawford et al., 2014).

[4] It has been shown that economic preferences and personality traits are stable during adolescence (Sutter et al., 2013; Harbaugh et al., 2002; Borghans et al., 2008). In particular, Sutter et al. (2013) find that economic preferences of children and adolescents are stable in the age group of 10 to 18 years old.

[5] This assumption is supported by the non-significance of the coefficient for grade in all our regressions with behavioural variables as the outcome of interest.

well as some of the students born in July 2000 in New South Wales. In order to control for a potential selection bias, we follow (Bedard and Dhuey, 2006) and instrument the observed relative age with the *predicted* relative age. A child is predicted to be relatively old if born just after the cut-off date, and relatively young if born just before the cut-off date. It is a strong predictor for a child's relative age in the classroom.

This instrumental variable regression gives an estimation of the local average treatment effect (LATE) of being relatively old in a cohort on the population of students complying with the cut-off date rule. Under the assumption that the position of the student's date of birth relative to the cut-off date is random, this estimation reflects the causal effect of the relative age. The available evidence suggest that this assumption is reasonable in Australia. Gans and Leigh (2009) have shown that there is no evidence for differences in birth rates in Australia before and after the school entry cut-off date.[6]

Hence, we estimate the parameters of equation (1) using the two-stage least squares procedure. Denoting ra_i^p as the predicted relative age given the child birth date and state cut-off date, the first-stage equation is defined as follows:

$$ra_{ij} = \theta_0 + \theta_1 ra_i^p + \theta_2 state_i + \mathbf{X}_i\theta_3 + \mathbf{S}_{ij}\theta_4 + \eta_{ij} \tag{2}$$

The standard errors are clustered at the school level, in order to account for the fact that the error term is not likely to be independent across observations within schools, e.g. students from the same school are exposed to the same environment which may influence the outcome variable (Angrist and Pischke, 2008), and also in order to account for experimental session effects.[7]

The estimation on the subsample of students born close to the cut-off date has the advantage of being a clear quasi-experiment: students were born very close in time and only small variations in their birthdate allocated them to one or the other side of the cut-off date. Restricting the estimation on this subsample has however a limitation: estimates will conflate any relative age effect with a potential effect of being in different grades. Given the different effects which can be reflected in the RDD estimate, Crawford et al. (2014) used two complementing empirical approaches to help disentangle the different possible effects. Similarly, we also run an analysis on an extended sample comprising youngest and oldests students in both grades 8 and 9. When estimating our model over the whole sample, we can include a dummy variable for grade and assess whether there is no grade effect as suggested by previous research (Sutter et al., 2013; Harbaugh et al.,

[6]This result is similar to an absence of break in the density of observations around a threshold which is a usual robustness check of the validity of the identification assumption underlying a regression discontinuity (Lee and Lemieux, 2010).

[7]Only one experimental session was conducted in each school (with the exception of two schools, where we divided the sample of participants and conducted two sessions).

2002). Here again, as for Crawford et al. (2014), the complementing approach does not identify stricly the age effect. When running the estimation on the whole sample, the coefficient of the relative age variable, will also incorporate a potential effect of absolute age. Calling β_1^{es} the coefficient of the relative age variable estimated on the extended sample, we have: $\beta_1^{es} = f(relative\ age, absolute\ age)$. For the same reasons as before, we expect preferences to be stable over one year of aging during school years. We therefore also expect that $\beta_1^{es} \approx \beta^{ra}$.

In the most general case, if grades or absolute age have an effect, it is not possible to uniquely identify the relative age effect. However, we can partially identify the relative age effect in the case where the absolute age effect and the grade effect go in the same direction. We call this the *maturity hypothesis*: if one more year of schooling and being one year older have an effect of preferences, this assumption posits that the effect would go in the same direction as it would represent a greater maturity from the student. Under the maturity hypothesis, our two parameters provide boundaries for the true relative age effect. Suppose for instance that grade and absolute age have a positive effect on a certain type of preferences, then $\beta_1^{qe} < \beta^{ra}$ because the oldest in the quasi-experimental setting have a lower grade. Conversely, $\beta_1^{es} > \beta^{ra}$ because the relatively old in the extended sample are also older in absolute age. As a consequence, $\beta_1^{qe} < \beta^{ra} < \beta_1^{es}$. If grade and absolute age have jointly a negative effect on the preference variables we would instead have $\beta_1^{qe} < \beta^{ra} < \beta_1^{es}$. As a consequence, even if grade or absolute age were to have an impact, our two estimates would likely provided boundaries for the true relative age effect.

3 Experimental design

3.1 Participant pool

The study was conducted in two Australian states, Queensland and New South Wales, where the cut-off date to start school is defined in different months. In Queensland, the experiments were conducted between March and November 2013, with a total of 547 high school students enrolled in grades 8 and 9, and born within a one-month window on each side of the cut-off date defined by the Queensland Department of Education. Until 2007, children in Queensland had to be 6 years old on or before the 31st of December in order to start school (grade 1) in that year's intake, starting usually at the end of January or beginning of February. Children born after the 31st of December would be eligible to start grade 1 in the next calendar year.[8] In New South Wales, we conducted

[8]Note that there was a change in the school entry cut-off date from 2008 in Queensland (the cut-off date switched from the 31st of December to the 30th of June). Students enrolled in 2013 in grades 6

the experiments between April and May 2014, with a total of 114 students enrolled in grades 8 and 9, and born either in July or August. Children in New South Wales are eligible to enter kindergarten, the first year of formal education in the state, if they turn 5 years old on or before the 31st of July.[9]

A total of 38 schools (32 in Queensland and 6 in New South Wales) accepted to take part in the experiment. Our sample in Queensland includes 19 schools in the Brisbane metropolitan area (capital of the state), 4 regional schools in other metropolitan areas, and 9 regional schools in remote or rural areas. Out of the 32 participating schools, 23 are public and 9 are private schools.[10] The sample in New South Wales includes public schools only, 5 located in the Sydney metropolitan area and 1 regional school.

An invitation to participate in the experiment was sent to all students enrolled in grades 8 and 9 and born in the months of December or January in Queensland, and July or August in New South Wales. In order to participate in the experiment, eligible students had to return an information letter briefly describing what their participation would involve, with the consent of their parents or caregivers for participation. Across the participating schools, an average of 47 percent of grade 8 and 42 percent of grade 9 students born in the months of interest took part in the experiment (difference not significant with the Student test, p=0.17). We do not observe significant differences in the participation rates between students born in December and January in Queensland, and July and August in New South Wales (see Appendix for details).

We excluded students who had been previously enrolled in other Australian states or countries with different school entry rules, since including these students would potentially bias our estimates for the relative age effect. There are a number of "non-compliers" in our sample with those students who had their school entry delayed (17%). This partial compliance before the cut-off dates warrants the use of an IV approach as in previous studies.[11]

and 7 (last two years of primary school) were subject to different school entry rules (as students enrolled in grade 6 in 2013 were the first cohort subject to the new school entry cut-off date). By focussing on students enrolled in grades 8 and 9 our sample is not affected by the change in cut-off date.

[9]Our choice of investigating relative age differences in economic preferences with adolescent students, in particular instead of younger children, was motivated by the fact that students in this age range had a longer experience of their relative (physical and cognitive) maturity in the classroom environment. A second reason for choosing this age range is the stability of preferences during adolescence (see Borghans et al. (2008), Sutter et al. (2013) and Harbaugh et al. (2002)). Furthermore, practical considerations also influenced our choice, as these are the first two years of high school in Queensland (unlike in New South Wales where high school starts with grade 7). We also expected these students to be mature enough to understand the experimental tasks and be able to stay focused during a 90-minute session.

[10]All public schools are co-educational while 6 private schools are single-sex and 3 co-educational.

[11]Further descriptive statistics are included in Appendix.

Table 1: Descriptive statistics

	Month of birth					
	December/July		January/August		Total	
	No.	%	No.	%	No.	%
Grade 8	182	55	177	54	359	54
Grade 9	151	45	151	46	302	46
Total	333	100	328	100	661	100
Female	155	47	166	51	321	49
Male	178	53	162	49	340	51
Total	333	100	328	100	661	100
Private school	71	21	65	20	136	21
Public school	262	79	263	80	525	79
Total	333	100	328	100	661	100
English grade						
A	38	13	45	16	83	14
B	135	46	107	38	242	42
C	96	32	99	35	195	34
D	22	7	28	10	50	9
E	5	2	3	1	8	1
Total	296	100	282	100	578	100
Math grade						
A	64	22	57	20	121	21
B	86	29	90	32	176	31
C	98	33	94	33	192	33
D	39	13	37	13	76	13
E	6	2	4	1	10	2
Total	293	100	282	100	575	100
Speak English at home						
No	32	10	28	9	60	9
Yes	296	90	298	91	594	91
Total	328	100	326	100	654	100
Mother's education						
No schooling	1	0	3	1	4	1
Primary school	21	6	15	5	36	5
High school	66	20	90	27	156	24
University degree	92	28	93	28	185	28
Doesn't know	153	46	127	39	280	42
Total	333	100	328	100	661	100
Father's education						
No schooling	3	1	4	1	7	1
Primary school	28	8	30	9	58	9
High school	64	19	85	26	149	23
University degree	90	27	73	22	163	25
Doesn't know	148	44	136	41	284	43
Total	333	100	328	100	661	100

As shown in Table 1, our sample counts with 328 students born in January or August and 333 students born in December or July. Fifty-four percent of the students in our sample are enrolled in grade 8, and 46 percent are enrolled in grade 9. Our sample is balanced in terms of gender. Seventy-nine percent of the students are enrolled in public schools, and 21 percent in private schools. We also have information on students' academic achievement, given by their grades in math and English, for 35 out of the 38 schools.[12] For the majority of the schools, the grades were obtained by students' self-report and also via the principal teacher of the school, whereas in a small number of schools, they were obtained uniquely by students' self-report. This is mainly due to the fact that some schools have very strict confidentiality rules. However, in the cases where we dispose from both information sources, we observe a very strong consistency between students' self-reported and school principals' grades. Therefore, self-reported information with respect to the grades is considered to be reliable. Many students also declared that they did not remember their grades. This, combined with the fact that we do not have information on the grades from the principal teacher for all the participating schools, explains why we do not have data on grades for all the students in the sample.

In line with the literature on the relative age effect, we observe a difference between students born before and after the cut-off date in the probability to repeat at least one grade (14 percent among those born in December or July, and 3 percent among those born in January or August). We do not observe any significant differences between students born before and after the cut-off date, in the academic results for both math and English (p-value with a t-test is 0.93 and 0.72, respectively).[13]

3.2 Experimental tasks

We elicited preferences for competitive environments and self-confidence as in Niederle and Vesterlund (2007), risk and ambiguity preferences as in Sutter et al. (2013) and Lejuez et al. (2002), and professional aspirations based on the work developed by Siegel (1957) and Page et al. (2007). We also collected socio-demographic variables through a post-experimental questionnaire. The experimental tasks were chosen taking into account their feasibility, with particular attention to the young subject pool, and also in order to follow the existing literature, facilitating the comparison of our results with previous findings.

All experimental sessions were conducted with the same experimenter (author in this study) plus either another author in the study or a research assistant. A teacher was always present during the session, so that the students would perceive the experiment

[12] We did not ask for the grades of the students in the first three schools where we conducted the experiment.

[13] We also present in Appendix A.4 the IV regression results for the impact of the relative age, instrumented by the month of birth, on academic scores.

as a formal activity. We visited each school only once during regular school hours in order to run the experiment. Almost all sessions were conducted in the morning (with one exception) and lasted between 70 and 90 minutes, depending on the number of participants. The instructions for all tasks were given out loud at the beginning of the session, following a written script. Examples were used to illustrate each of the tasks and participants were encouraged to ask questions in order to ensure a good understanding of the tasks and the associated payoff structure. Subsequently, the students performed the experimental tasks individually, on a website specifically designed for our experiment. Once the experiment started, all questions were answered privately. The experimental webpage displayed screens with a brief summary of the instructions before the participants started to perform each of the tasks.

The experiment was incentivised with real monetary payments. Each student received a participation fee of 7 Australian dollars and could earn up to 14 Australian dollars, depending on his or her choices and performance at the tasks. The upper bound for the gains was defined taking into account the age of our participants, such that their earnings would correspond approximately to the average weekly pocket money available to young students. Task 4, measuring aspirations was the only one not to be incentivised. One out of the three incentivised tasks was selected randomly at the end of the session to determine the participant's final gains. These were distributed privately to the participants at the end of the session.

3.2.1 Preferences for competitive environments

Our first task builds on Niederle and Vesterlund (2007),[14] in which the participants have to solve a simple math task, consisting of adding sets of five one-digit or two-digit numbers, under different payment schemes. The task comprises five stages (in fixed order). In each of the first three stages, participants are asked to successively add up sets of five numbers (as illustrated in Figure A.2 in Appendix A.3) during three minutes. Once they submit an answer, a different set of five numbers is displayed on the screen. Participants are not given any feedback on whether the previous answer was right or wrong.

In the first stage, participants face a piece-rate compensation scheme and receive 20 cents for each correct answer. In the second stage, participants are exposed to a tournament compensation scheme. They receive 60 cents per correct answer if they are ranked in the top third in terms of number of correct answers in stage 2. If they are

[14]The main difference with our experiment is that, in Niederle and Vesterlund (2007) participants are matched within mixed groups of four subjects and the participant with the highest number of correct answers in each group wins the tournament. In our setting, there is no matching of the participants and the top third of participants in the classroom win the tournament.

ranked below the top third, their payoff is zero. These two stages provide participants with experience in both compensation schemes.

In the third stage, participants have to choose which compensation scheme will be applied to their future performance. They are then given another three minutes to solve the calculations. The performance of a participant choosing the competitive compensation scheme, is evaluated relative to the performance of all participants observed in the previous stage (stage 2). The performance of a participant who enters a tournament is compared to the performance of participants who also performed under a tournament compensation scheme.

In the fourth stage of the task, participants do not have to perform any calculations, but choose the compensation scheme for their past piece-rate performance (in stage 1). They are first informed about their number of correct answers in stage 1 (but not about their rank), and have to choose between the piece-rate and the tournament compensation scheme.[15] This stage allows to test whether the heterogeneity between different groups in preferences for competition (assessed in stage 3), can be explained solely by differences in taste for performing under a competitive environment, or whether it can also be explained by other factors, such as over-confidence and risk aversion, as the competition makes the payoff more uncertain (Niederle and Vesterlund, 2010).

In the final stage of the task, participants are asked to guess their relative performance in stages 1 and 2. The elicitation of beliefs for relative performance is not incentivised, in order to ensure that participants have no interest in behaving strategically in the first stages, for instance by performing very poorly.

3.2.2 Risk and ambiguity attitudes

Ordered choices lists We use ordered choices lists as in Sutter et al. (2013) in order to elicit risk and ambiguity attitudes. In each list, participants have to choose twenty times between drawing a ball from a bag (filled with white and orange balls) and an increasing sure amount of money, as shown in Appendix A.3.2. If participants choose to draw a ball from the bag, they earn 7 dollars if the drawn ball is white, and 0 is the drawn ball is orange. At the end of the session, only one out of the twenty decision rows is randomly selected by the experimental software to determine participants' earnings. If participants chose to draw a ball from the bag in the selected row, they will draw a ball from the bag themselves and be rewarded according to the colour of the ball drawn. If they chose the sure option, the amount of money offered in the selected row will be paid to the participant.

[15] If a participant chooses to submit his or her piece-rate performance to a tournament, his or her score will be compared to the score in stage 2 of the other participants in the session.

Participants are presented with 2 different bags, named A and B. Bag A, represents the risky prospect and is filled with 6 white and 6 orange balls in front of the participants. Bag B, represents the ambiguity prospect and is filled before the experimental session. The participants were told that the bag contained 12 balls, but they were not informed about the proportion of each colour (several examples are used during the instructions for the potential proportions of white and orange balls in the bag, in order to avoid influencing their expectations on their probability of drawing a white ball). Based on participants' choices between the sure amount of money and the random draw, we calculate the parameters for risk and ambiguity attitudes as in Sutter et al. (2013) (a detailed description is provided in Appendix A.3.2).

In order to ensure a good understanding of the task, participants were given different examples of completed choice lists (corresponding to different risk attitudes) and we explained what would be the outcome in terms of payoffs if a given decision row was randomly chosen out of the 20 rows of a filled list. The different examples were presented in random order and we did not mention whether a given choice list represented the preferences of a more or less risk seeking individual. We also explicitly guided the participants to make consistent choices. We explained to the participants that once they preferred the sure amount of money offered over drawing a ball from the bag, they should also prefer the sure amount of money in all the subsequent rows. Accordingly, as in Andersen et al. (2006), our experimental software limited the possibility to make inconsistent choices (i.e. participants could not have more than one switching point between the risky and the sure option).

Balloon Analogue Risk Task We use the BART task (Lejuez et al., 2002) (commonly used in psychology studies) as a second measure of risk attitudes. In this task, a simulated balloon with a balloon pump is presented to the participants on the computer screen (see Appendix A.3.3). Each click on the pump inflates the balloon and increases participant's earnings by 5 cents. The balloon will explode at a random pump, leading to a payoff of zero. At any time before its explosion, participants can decide to stop pumping the balloon and collect the accumulated earnings. Participants were not given information regarding the probability of the explosion of the balloon, but they were told that the balloon can explode at any point, between the first click all the way up to the maximum of 140 clicks, at which point the balloon will fill the entire computer screen. After each balloon explosion or money collection, participants were presented with a new balloon, until a total of five balloons had been completed. Our measure of risk attitude is given by the total number of clicks.

The BART can therefore be considered as a way to elicit risk attitude in a situation where probabilities are subjective. It is a task which is easy to understand for young participants and its simplicity, makes it a natural complement to other more formal economic tasks to elicit risk attitude (see Charness et al. (2013) for an extended discussion). The answers to the BART have been found to correlate with real life risk behaviour (Lejuez et al., 2003, 2007).

3.2.3 Professional aspirations

Aspiration levels can be considered as reference points in a prospect theory framework Page et al. (2007). Reference points split the space of outcomes in a way such that the marginal utility of gains is larger below the aspiration level than above (Kahneman and Tversky, 1979; Diecidue and Van de Ven, 2008). We rely on this specific distortion of the utility space to elicit aspiration levels, in way related to the approach suggested by Siegel (1957).

We measure participants' aspirations by their ranking of professional careers on a preference scale (see Appendix A.3.4). In this task, the participants are presented with three different occupations. Each of the three occupations is associated with a different social status and a different rank in the official classification of occupations. Two of the three occupations are placed at the upper and lower end of a normalised scale of satisfaction. The occupation associated with the lowest social status is assigned to the bottom of the scale, and the one associated with the highest social status is assigned to the top of the scale. The task consists in ranking the third occupation, associated with an intermediate rank in the classification of occupation and social status, on this preference scale.

Our measure of aspirations is given by the distance between the lowest ranked profession (fixed on the scale) and the third occupation, ranked by the participants. If the intermediate profession provides a satisfaction very close to the least attractive profession, this indicates high aspirations: only the high ranked profession appears satisfactory to the participant. In contrast, if the intermediate occupation provides a satisfaction close to the occupation associated with the highest rank and social status, this reveals low aspirations. We use occupations from different professional areas (scientific, artistic, educational) to capture aspirations over a diverse range of professional careers.

4 Results

We focus on the estimates obtained with the IV method, net of the potential selection bias, for interpretation of the results and report the estimates obtained with the OLS procedure for comparative purposes in Appendix B. The estimates from the first stage of the IV regression, also in Appendix A.5, show that the predicted relative age (ra^p), given by whether a child was born before or after the cut-off date, is a strong predictor for the observed relative age.[16]

Result 1: *We have limited evidence of a relative age effect on preferences for competitive environments.*

When investigating the possible role of relative age differences in preferences for competition, we do not find evidence of any causal link between relative age and performance. However, we find in some specifications some evidence of a relative age effect on preferences for competition.

Result 1.1: *Relatively old students do not outperform relatively young students in the tournament.*

In Table 2, we show the estimation results with the number of correct answers in stage 1 (piece-rate), the number of correct answers in stage 2 (tournament) and the progression in performance between the two stages as dependent variables. The results indicate that relatively old students do not perform better under the piece-rate or the tournament.

In the specifications with additional control variables, we find that students in grade 9 perform better than students in grade 8.

They obtain on average 0.69 more correct answers in the piece-rate (specification (1b)) and 0.86 more correct answers in the tournament compensation scheme (specification (2b)). The estimate for grade 9 combines the effect of a one year difference in formal education and in chronological age on performance. As in Niederle and Vesterlund (2007) and Sutter and Rützler (2010), we do not find any gender differences in performance.[17]

We also find that, as expected, students with higher math scores perform better. An increase in one standard deviation from the mean in math scores increases the number of correct answers by 1.1 in the piece-rate (stage 1) and by 1.25 in the tournament payment scheme (stage 2).

[16] There are slight variations in sample sizes between the different regression specifications. One of the reasons is that, as mentioned in section 3, we do not have data on students' academic grades for all participating schools. A second reason is that as the tasks were performed online, we sporadically randomly lost information on participants' choices due to internet connection issues during the task. As the data is missing at random, it does not affect the consistency of the estimations (Heitjan and Basu, 1996). Finally, in the ordered choices list task, we excluded 5% of participants who made inconsistent choices (by chosing a first ordered stochastically dominated choice in the bottom row) as this might be an indication that they did not understand the task.

[17] Other studies provide evidence that competition increases the gender gap in performance. Gneezy et al. (2003) conduct an experiment with Israeli engineering students, who had to solve mazes under different incentive schemes. They find that competition increases by factor three the gender gap in performance compared to the non-competitive setting. They also find that the gender gap in performance is not due to the fact that women do not perform well under competitive environments generally, but that they do not perform well when competing in mixed-gender groups.

Table 2: IV regressions on performance

	Quasi-experimental sample						Extended sample					
	Piece-rate		Tournament		Difference		Piece-rate		Tournament		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
Relative age	-0.239	0.051	-0.922	-0.688	0.676	0.783	0.686	0.401	0.135	-0.134	0.521	0.494
	(1.036)	(1.130)	(1.099)	(1.178)	(0.664)	(0.705)	(0.587)	(0.588)	(0.664)	(0.691)	(0.463)	(0.507)
QLD	-0.611	-0.482	-1.297	-1.484**	0.683*	1.032***	-0.876	-0.575	-1.164	-0.944*	0.305*	0.377
	(0.885)	(0.741)	(0.898)	(0.679)	(0.379)	(0.393)	(0.665)	(0.511)	(0.715)	(0.490)	(0.183)	(0.263)
Grade 9								0.691**		0.864***		-0.200
								(0.297)		(0.290)		(0.213)
Male		0.625		0.260		0.339		0.281		0.102		0.163
		(0.526)		(0.535)		(0.291)		(0.284)		(0.343)		(0.236)
Constant	6.348***	6.979***	8.877***	10.698***	-2.521***	-3.793***	6.057***	0.776	7.698***	1.291	-1.633***	-0.248
	(1.789)	(1.743)	(1.886)	(1.804)	(0.881)	(1.021)	(1.304)	(3.038)	(1.459)	(3.136)	(0.518)	(1.899)
N	283	252	283	252	282	251	534	463	532	461	529	458
Add. controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 1.2: *We find some evidence that relatively old students may be more likely to choose to perfom in a competitive environment.*

In stage 3, participants were first asked to choose between one of the payment methods, piece-rate or tournament, that would be applied to their future performance. In our sample, 35.11 percent of the participants choose to enter the tournament.

Table 3 presents the results obtained with the probit regressions where the dependent variable is equal to 1 if the participant decides to enter the tournament and 0 if he or she chooses the piece-rate. When looking at the quasi-experimental subsample, we find that relatively old students are more likely to choose the competitive setting for their payoffs. Splitting the sample by gender reveals that this effect is entirely concentrated on male students. For them, being the oldest grade 8 make them ....% more likely to chose a competitive environment than the youngest in grade 9 when controling for their other individual characteristics.

Noticeably, this effect vanishes when we run the estimation on our whole sample. Theoretically such a discrepancy between these two estimates could come from the fact that the effect found on the quasi-experimental subsample is driven by a grade effect which is controlled for in the whole sample. However this is not the case as the grade effect is never close to conventional levels of significance.

Looking further at the cause of this difference, we found that it comes from different results between grade 8 and grade 9. In grade 8, relatively old people are markedly more likely to chose the competitive setting while there is no such effect in grade 9. The estimation on the whole sample therefore dilutes the effect found on the quasi-experimental subsample.

As in Niederle and Vesterlund (2007) we find strong evidence for gender differences in preferences for competition. Male students are more likely to choose the tournament compared to female students, even in the absence of differences in performance and controlling for math scores and performance in stage 2. Specification (1b) in Table 3 shows that, controlling for beliefs for performance, female students have on average a 38.6 percentage point lower probability of choosing the tournament (significant at 1 percent level), compared to male students. This very large effect size is in line with the results in Niederle and Vesterlund (2007), who find that females are 38 percentage points less likely to enter the tournament than males. Moreover, the probability of entering the tournament decreases with the number of participants in the session, although the tournament rule was independent of the number of participants (participants had to be ranked in the top third in order to win the tournament). The consistency between these additional results and previous studies provide support to our experimental design and its ability to detect differences in preference for competition.

Table 3: IV probit regressions for the choice of compensation scheme (stage 3)

	Quasi-experimental sample				Extended sample			
	All		Male	Female	All		Male	Female
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Relative age	0.533^{+}	0.776**	1.484***	0.390	0.073	0.100	0.206	0.098
	(0.310)	(0.327)	(0.562)	(0.486)	(0.228)	(0.244)	(0.440)	(0.332)
QLD	0.070	0.064	-0.151	0.209	-0.123	-0.188	-0.391^{+}	0.140
	(0.190)	(0.232)	(0.276)	(0.364)	(0.120)	(0.159)	(0.230)	(0.244)
Grade 9						-0.029	-0.063	-0.021
						(0.133)	(0.159)	(0.238)
Male		0.205				0.386***		
		(0.187)				(0.150)		
Tournament performance		0.046	0.051	0.044		0.043^{+}	0.041	0.045
		(0.032)	(0.032)	(0.052)		(0.023)	(0.030)	(0.035)
No. of participants		-0.018^{+}	-0.007	-0.028^{+}		-0.016***	-0.009	-0.031***
		(0.010)	(0.011)	(0.017)		(0.006)	(0.008)	(0.008)
Rank guessed at tournament		-0.019***	-0.019***	-0.021***		-0.021***	-0.020***	-0.023***
		(0.004)	(0.005)	(0.007)		(0.003)	(0.005)	(0.006)
N	279	248	126	122	529	455	242	213
Add. controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Average marginal effects; Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 1.3: *Relatively old students do not display an overconfidence in regard to their relative ability among other students.*

In stage 4, participants choose between a competitive and non competitive setting for the reward of their past performance. It is very similar to the choice to enter a competition, though they do not have to perform in a tournament. A tendency to opt for the competitive setting here may therefore reflect a propensity to be overconfident.[18] Overall, we observe that 32.28 percent of the participants choose to submit their piece-rate performance to a tournament, which is a very similar proportion to those who opted for the tournament in stage 3 (35.11 percent). When looking at the determinants for this decision (Table 4), we do not find any significant relative age effect on the decision to submit the past performance to a competitive payoff scheme. This is true both in the quasi-experimental sub-sample and in the overall sample.

We however find a significant gender difference in overconfidence. Controlling for participants' beliefs for their rank in the piece-rate in specification, male participants are 47.1 percentage points more likely to submit their performance to the tournament than female participants (specification (1b)).

[18] Or greater preferences for risk which we investigate in the other tasks.

Table 4: IV probit for the submission of the piece rate to a tournament (stage 4)

	Quasi-experimental sample				Extended sample			
	All		Male	Female	All		Male	Female
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Relative age	0.360	0.581	0.631	0.441	0.167	0.154	0.113	0.212
	(0.352)	(0.417)	(0.746)	(0.587)	(0.225)	(0.275)	(0.489)	(0.360)
QLD	-0.092	0.131	0.228	-0.020	-0.138	-0.121	-0.005	-0.223
	(0.248)	(0.255)	(0.264)	(0.420)	(0.124)	(0.174)	(0.196)	(0.332)
Grade 9						0.101	0.105	0.175
						(0.119)	(0.175)	(0.174)
Male		0.299				0.471***		
		(0.227)				(0.149)		
Piece rate performance		0.017	0.062	-0.032		0.019	0.044^{+}	-0.018
		(0.027)	(0.039)	(0.052)		(0.019)	(0.023)	(0.037)
No. of participants		-0.033***	-0.034**	-0.038**		-0.023***	-0.022***	-0.030**
		(0.011)	(0.014)	(0.016)		(0.006)	(0.008)	(0.012)
Rank guessed at piece-rate		-0.022***	-0.026***	-0.017**		-0.018***	-0.020***	-0.018***
		(0.005)	(0.008)	(0.008)		(0.003)	(0.005)	(0.006)
N	281	249	126	123	533	459	241	218
Add. controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Average marginal effects; Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To complement this choice task, participants were also asked to report their guesses for their rank, for both stage 1 (piece-rate) and stage 2 (tournament).[19] In tables 5 and 6 we show the regression results for participants' expectations for their rank in stages 1 and 2, respectively.

We do not find any relative age difference in participants' expectations for their relative performance (specifications (1a) and (1b)).

The results also indicate that students in grade 9 expect to be ranked higher than students in grade 8, both in the piece-rate and tournament. Moreover, we find evidence for a large gender gap in beliefs for relative performance, even when controlling for math scores and performance. Male participants expect to be ranked 7.03 percentiles higher than female participants under the piece-rate, and 7.89 under the tournament (in specification (2b)). Finally, we observe that participants who have better math scores and performed better at the task expect to be ranked higher.

Overall, we do not find evidence that relatively old students are more overconfident than relatively young ones. Neither their choices nor the elicitation of their beliefs on how their performance fare relative to others suggest such differences. However, we find evidence of significant gender differences. Male students appear more confident about their performance than female students. This result is consistent with the previous results from Niederle and Vesterlund (2007) and Sutter and Rützler (2010).

[19]The variable *expected rank* is expressed in percentile rank. It is calculated by dividing the guessed rank of the participant by the total number of participants in the session and then multiplying this fraction by 100.

Table 5: IV regressions for the expected rank in the piece-rate (stage 1)

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	1.212	6.968	13.491**	-4.641	-0.827	5.533
	(8.125)	(7.270)	(6.836)	(4.215)	(4.139)	(4.773)
QLD	-0.071	-2.534	-1.603	-0.024	-5.301	-4.908
	(2.872)	(4.718)	(4.882)	(2.927)	(3.798)	(3.813)
Grade 9					-6.337***	-5.853***
					(1.959)	(1.974)
Male		-8.371***	6.025		-7.033***	5.012
		(2.887)	(13.847)		(1.965)	(7.436)
Relative age x Male			-17.875			-15.463
			(15.674)			(9.674)
Piece rate performance		-2.704***	-2.802***		-2.966***	-3.040***
		(0.477)	(0.501)		(0.316)	(0.318)
No. of participants		0.020	-0.024		-0.211	-0.240+
		(0.244)	(0.245)		(0.140)	(0.138)
Constant	51.671***	75.999***	71.653***	54.895***	144.621***	136.290***
	(8.387)	(15.766)	(15.720)	(5.851)	(21.128)	(22.333)
N	284	252	252	537	463	463
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area.
Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: IV regressions for the expected rank in the tournament (stage 2)

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	2.266	8.281	10.164	-3.634	1.117	1.836
	(7.091)	(6.866)	(7.502)	(4.441)	(4.287)	(5.156)
QLD	-0.329	-0.922	-0.663	2.446	-0.416	-0.373
	(2.893)	(4.588)	(4.679)	(2.952)	(2.979)	(3.035)
Grade 9					-5.589***	-5.530***
					(1.955)	(1.985)
Male		-12.250***	-8.092		-7.886***	-6.523
		(2.953)	(13.930)		(2.214)	(7.098)
Relative age x Male			-5.183			-1.752
			(16.251)			(9.489)
Tournament performance		-1.444**	-1.464**		-1.801***	-1.807***
		(0.705)	(0.709)		(0.459)	(0.455)
No. of participants		0.284	0.271		-0.023	-0.026
		(0.192)	(0.193)		(0.107)	(0.108)
Constant	51.399***	63.841***	62.585***	49.446***	119.944***	118.968***
	(8.235)	(17.315)	(17.668)	(6.384)	(20.916)	(22.063)
N	284	252	252	537	461	461
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area.
Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 2: *We only find limited evidence of a possible relative age effect on attitudes toward risk and ambiguity for males.*

Result 2.1: *Relatively old students do not have different attitudes toward risk than relatively young students.*

In Table 7, we present the regression results for risk preferences, using the ordered choices list task[20]

In Table 7, we show that the relative age does not seem to impact risk attitudes. Likewise, an additional year of formal education does not seem to significantly impact risk preferences. We also find that males are more risk seeking than females. Being male, as opposed to being female decreases the risk aversion parameter by 0.07. This result is consistent with Sutter et al. (2013), who find that males are less risk averse (by 0.07) than female students. Moreover, an increase in one standard deviation from the mean in math scores decreases the risk aversion parameter by 0.05. This result is in line with the finding in Burks et al. (2009) that individuals with higher IQ have lower levels of risk aversion, when measured using binary choices between safe amounts of money and a lottery.

We also did not find any differences for the BART task. We follow Lejuez et al. (2002) and measure risk aversion by the average adjusted number of pumps (which is given by the average number of pumps excluding balloons that exploded, given that the number of pumps is constrained on balloons that exploded).[21]

The regression results in Table 8 show that there is no evidence of a relative age effect on risk aversion. Moreover, we do not find evidence of gender differences when considering as a measure of risk aversion the adjusted average number of pumps in the BART.

[20] We excluded from the regression analysis participants who never switch from the risk or ambiguous prospect to the sure option (approximately 5 percent of the sample), as this might be an indication that they did not understand the task.

[21] We excluded from the sample 8 percent of the participants for whom we do not have information on each of the five balloons, due to internet connection issues during the task. It is very unlikely to change the results as these events were randomly distributed across participants.

Table 7: IV regressions on risk aversion

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	0.056	0.050	0.087	0.001	0.022	0.025
	(0.062)	(0.071)	(0.084)	(0.039)	(0.045)	(0.059)
QLD	0.025	0.042	0.048	0.019	0.014	0.014
	(0.032)	(0.045)	(0.046)	(0.023)	(0.032)	(0.031)
Grade 9					-0.035	-0.035
					(0.032)	(0.033)
Male		-0.073**	0.006		-0.074***	-0.069
		(0.034)	(0.161)		(0.024)	(0.097)
Relative age x Male			-0.099			-0.006
			(0.194)			(0.125)
N	268	238	238	505	438	438
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: IV regressions on the BART

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	-12.395	-11.362	-19.782	-17.089	-18.185	0.307
	(28.657)	(27.964)	(27.438)	(17.055)	(17.846)	(18.401)
QLD	-0.210	-8.219	-9.616	-2.045	-8.599	-6.851
	(15.712)	(19.175)	(17.983)	(14.995)	(18.808)	(19.284)
Grade 9				0.878	0.044	1.302
				(7.086)	(7.145)	(7.504)
Male		5.764	-11.840		11.204	45.309+
		(11.012)	(45.632)		(9.720)	(26.742)
Relative age x Male			22.276			-44.200
			(50.430)			(28.794)
N	259	242	242	488	450	450
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 2.2: *Relatively old students do not have different attitudes toward ambiguity* *than relatively young students.*

In Table 9 we present the regression results on ambiguity aversion. As for risk attitudes with the ordered choices list task, we do not find evidence that the relative age impacts ambiguity preferences. As in Sutter et al. (2013), we also do not find evidence for gender differences.

Table 9: IV regressions on ambiguity aversion

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	-0.013	0.021	-0.077	-0.013	-0.012	-0.095
	(0.059)	(0.064)	(0.069)	(0.040)	(0.045)	(0.077)
QLD	-0.010	-0.029	-0.045	0.015	0.012	0.003
	(0.043)	(0.043)	(0.040)	(0.039)	(0.042)	(0.040)
Grade 9					-0.018	-0.025
					(0.019)	(0.018)
Male		-0.005	-0.212		-0.016	-0.167^{+}
		(0.039)	(0.136)		(0.030)	(0.101)
Relative age x Male			0.262^{+}			0.196
			(0.157)			(0.123)
N	264	237	237	502	438	438
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 3: *We do not find evidence of a relative age effect on professional aspirations.*

In Table 10 we show the regression results for the whole sample on the average rank across the four occupations. The results indicate that relative age is not an important predictor of aspirations.

Table 10: IV regressions on professional aspirations

	Average rank for the 4 occupations					
	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	0.560	-2.540	-6.977^{+}	-0.044	-1.415	-2.412
	(3.268)	(3.103)	(3.948)	(1.884)	(2.030)	(3.121)
QLD	-1.336	-0.699	-1.415	-2.768**	-2.243^{+}	-2.339
	(1.260)	(1.607)	(1.874)	(1.306)	(1.328)	(1.439)
Grade 9				2.050	2.394^{+}	2.313
				(1.585)	(1.431)	(1.479)
Male		-0.147	-10.022		-0.044	-1.943
		(1.952)	(9.241)		(1.294)	(4.670)
Relative age x Male			12.322			2.434
			(11.724)			(6.400)
N	279	250	250	524	461	461
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusion

This study investigated whether the persistent relative age gap in academic outcomes could be partly explained by relative age differences in economic preferences relevant for decision making in education.

We conducted the experiment in 38 high schools, with a sample of 661 adolescent students aged between 13 and 14 years old. All our participants were born within a two-month window of the cut-off date defined to start school.

To the best of our knowledge, it is the first investigation of the RAE on economic preferences. Even though we looked at a wide range of behavioural traits, we did not find substantial evidence of an effect of relative age on these. Given the large number of tests conducted in this study, this single result should be taken with caution when trying to interpret it.

Taken together, our results suggest that relative position in a peer group may primarily have an effect on confidence about ability at school as found by (Crawford et al., 2014). But that it the effect of relative age may not extend to broader economic preferences and psychological traits such as self-confidence, risk attitude and professional aspiration levels. However, clearly our study does not exhaust the question of the effect of relative position among peers on the formation of behavioural traits. In the specific case of the relative age effect, it would be interesting to test whether we observe difference in self-control and time preferences between relatively young and relatively old children. For instance, Bettinger and Slonim (2007) find that more mature children are more patient. Hence, testing whether relatively young children are less patient than their older counterparts could provide important insights on the type of pedagogical solutions which could be effective in attenuating the relative age effect in educational trajectories.

Overall, the main finding of our study is that relative age differences in academic outcomes found throughout the educational trajectory and in later outcomes are unlikely to be driven by systematic differences in behavioural traits such as self-confidence, preference for competition, risk attitude and professional aspirations. Given the importance of the relative age effect, and the existing evidence that relative age affects students' self perception, this result invites to study further the psychological mechanisms underlying the emergence of differences in outcomes between relatively old and relatively young students at school.

References

Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4):383–405.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Battaglini, M., Bénabou, R., and Tirole, J. (2005). Self-control in peer groups. *Journal of Economic Theory*, 123(2):105–134.

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

Bedard, K. and Dhuey, E. (2007). Is september better than January? The effect of school entry age laws on skill accumulation. Mimeo.

Bettinger, E. and Slonim, R. (2007). Patience among children. *Journal of Public Economics*, 91(1):343–363.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *The Journal of Human Resources*, 43(4):972–1059.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19):7745–7750.

Charness, G., Gneezy, U., and Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, 87:43–51.

Crawford, C., Dearden, L., and Greaves, E. (2014). The drivers of month-of-birth differences in children's cognitive and non-cognitive skills. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

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

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

Diecidue, E. and Van de Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility*. *International Economic Review*, 49(2):683–700.

Du, Q., Gao, H., and Levi, M. D. (2012). The relative-age effect and career success: Evidence from corporate CEOs. *Economics Letters*, 117(3):660–662.

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

Fenzel, L. M. (1992). The effect of relative age on self-esteem, role strain, GPA, and anxiety. *The Journal of Early Adolescence*, 12(3):253–266.

Fredriksson, P. and Öckert, B. (2006). Is early learning really more productive? the effect of school starting age on school and labor market performance. Working Paper, IFAU-Institute for Labour Market Policy Evaluation.

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

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Goodman, R., Gledhill, J., and Ford, T. (2003). Child psychiatric disorder and relative age within school year: Cross sectional survey of large population sample. *British Medical Journal*, 327(7413):472–475.

Grenet, J. (2009). Academic performance, educational trajectories and the persistence of date of birth effects: Evidence from France. Unpublished manuscript. London School of Economics.

Harbaugh, W. T., Krause, K., and Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, 5(1):53–84.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4):1009–1055.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the national Academy of Sciences*, 104(33):13250–13255.

Heitjan, D. F. and Basu, S. (1996). Distinguishing "missing at random" and "missing completely at random". *The American Statistician*, 50(3):207–213.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, (47):263–291.

Koch, A., Nafziger, J., and Nielsen, H. S. (2014). Behavioral economics of education. *Journal of Economic Behavior and Organization, forthcoming.*

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

Lejuez, C., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., and Gwadz, M. (2007). Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART–Y) in the assessment of risk-taking behavior among inner-city adolescents. *Journal of Clinical Child and Adolescent Psychology*, 36(1):106–111.

Lejuez, C., Aklin, W. M., Zvolensky, M. J., and Pedulla, C. M. (2003). Evaluation of the balloon analogue risk task (bart) as a predictor of adolescent real-world risk-taking behaviours. *Journal of adolescence*, 26(4):475–479.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., and Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). *Journal of Experimental Psychology: Applied*, 8(2):75–84.

Manski, C. F. (2000). Economic analysis of social interactions. *The Journal of Economic Perspectives*, 14(3):115–136.

McEwan, P. J. and Shapiro, J. S. (2008). The benefits of delayed primary school enrollment discontinuity estimates using exact birth dates. *Journal of Human Resources*, 43(1):1–29.

Morrow, R. L., Garland, E. J., Wright, J. M., Maclure, M., Taylor, S., and Dormuth, C. R. (2012). Influence of relative age on diagnosis and treatment of attention-deficit/hyperactivity disorder in children. *Canadian Medical Association Journal*, 184(7):755–762.

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

Müller, D. and Page, L. (2015). Political selection and the relative age effect. *QuBE Working Papers.*

Murphy, R. and Weinhardt, F. (2014). Top of the class: The importance of ordinal rank. CESifo Group Munich, Working paper no. 4815.

Musch, J. and Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review*, 21(2):147–167.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *The Journal of Economic Perspectives*, 24(2):129–144.

Page, L., Garboua, L. L., and Montmarquette, C. (2007). Aspiration levels and educational choices: An experimental study. *Economics of Education Review*, 26(6):747–757.

Puhani, P. and Weber, A. (2007). Does the early bird catch the worm? *Empirical Economics*, 32:359–386.

Siegel, S. (1957). Level of aspiration and decision making. *Psychological review*, 64(4):253–262.

Smith, J. (2009). Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement. *The BE Journal of Economic Analysis and Policy*, 9(1).

Sutter, M., Kocher, M. G., Glatzle-Rutzler, D., and Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *The American Economic Review*, 103(1):510–531.

Sutter, M. and Rützler, D. (2010). Gender differences in competition emerge early in life. Mimeo, University of Innsbruck.

Thompson, A. H., Barnsley, R. H., and Battle, J. (2004). The relative age effect and the development of self-esteem. *Educational Research*, 46(3):313–320.

Thompson, A. H., Barnsley, R. H., and Dyck, R. J. (1999). A new factor in youth suicide: The relative age effect. *Canadian Journal of Psychiatry*, 44:82–85.

A Appendix

A.1 Empirical strategy

In this section, we discuss the general case, that is when the impact of absolute age and grade cannot be assumed equal to zero.

We consider the following model for the relationship between performance and chronological age, schooling and relative age:

$$y_i = \beta_0 + \beta_1 age_i + \beta_2 grade_i + \beta_3 ra_i + \beta_4 state_s + \mathbf{X}_i\beta_5 + \mathbf{X}_j\beta_6 + \epsilon_i \quad (3)$$

We estimate jointly $\beta_1 + \beta_3$ and $\beta_1 + \beta_2$ by estimating the following equation:

$$y_i = \gamma_0 + \gamma_1 grade_i + \gamma_2 ra_i + \gamma_3 state_s + \mathbf{X}_i\gamma_4 + \mathbf{X}_j\gamma_5 + \nu_i \quad (4)$$

By estimating Equation (2) we obtain estimates of the following parameters:

$$\gamma_1 = \beta_1 + \beta_2 \quad \text{(5a)}$$

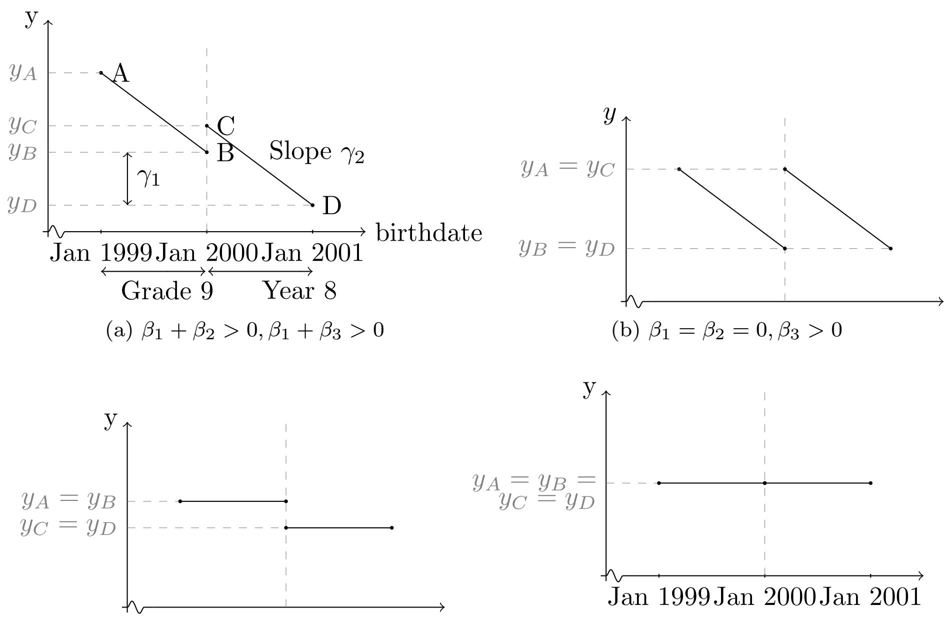
$$\gamma_2 = \beta_1 + \beta_3 \quad \text{(5b)}$$

In the specific case where there is no effect on outcomes (y_i) of a one year difference in chronological age or in schooling, γ_1 will be equal to zero, and therefore $\beta_1 + \beta_2$ will also be equal to zero (or $\beta_1 = \beta_2 = 0$ under the maturity assumption). Consequently, the relative age effect (β_3) is identified and given by γ_2 in Equation (2). This scenario is represented graphically in Figure A.1b.

Alternatively, if the combined effect of the chronological age and schooling is significantly different from zero ($\gamma_1 \neq 0$ in Equation (2)), then γ_2 is the upper bound of the relative age effect (β_3), averaged across school cohorts.

If $\beta_1 + \beta_3 = 0$ (or $\gamma_2 = 0$) and $\beta_1 + \beta_2 \neq 0$ (or $\gamma_1 \neq 0$), only grade has an impact on outcomes. This case corresponds to Figure A.1c.

Finally, if $\beta_1 + \beta_3 = 0$ (or $\gamma_2 = 0$) and $\beta_1 + \beta_2 = 0$ (or $\gamma_1 = 0$), none of the variables considered (age, grade and relative age) affects preferences. This situation is represented in Figure A.1d.



(a) $\beta_1 + \beta_2 > 0, \beta_1 + \beta_3 > 0$

(b) $\beta_1 = \beta_2 = 0, \beta_3 > 0$

(c) $\beta_1 = \beta_3 = 0, \beta_2 > 0$

(d) $\beta_1 = \beta_2 = \beta_3 = 0$

Figure A.1: Identification strategy

A.2 Descriptive statistics

Table A.1: Distribution of participants bewteen compliers and non-compliers, by month of birth

	Month of birth					
	July/Dec		August/Jan		Total	
	No.	%	No.	%	No.	%
Compliers	136	41	267	81	403	61
Non-compliers						
Redshirted	122	37	3	1	125	19
Repeaters	46	14	9	3	55	8
Grade advanced	0	0	5	2	5	1
Enroled in other states or countries	67	20	55	17	122	18

Figure A.3: List 1

1. draw from bag A	☒	or	☐	\$0.35 for sure
2. draw from bag A	☒	or	☐	\$0.70 for sure
3. draw from bag A	☒	or	☐	\$1.05 for sure
4. draw from bag A	☒	or	☐	\$1.35 for sure
5. draw from bag A	☒	or	☐	\$1.70 for sure
etc...				
16. draw from bag A	☒	or	☐	\$5.60 for sure
17. draw from bag A	☒	or	☐	\$5.95 for sure
18. draw from bag A	☐	or	☒	\$6.30 for sure
19. draw from bag A	☐	or	☒	\$6.65 for sure
20. draw from bag A	☐	or	☒	\$7.00 for sure

A.3 Experimental tasks

A.3.1 Task 1: Preferences for competitive environments

Figure A.2: Illustration of the math task

8	16	10	12	7	

A.3.2 Task 2: Ordered choices lists

As in Sutter et al. (2013), we denote r the risk attitude, which is obtained as follows:

$$r = 1 - CE_r/\pi$$

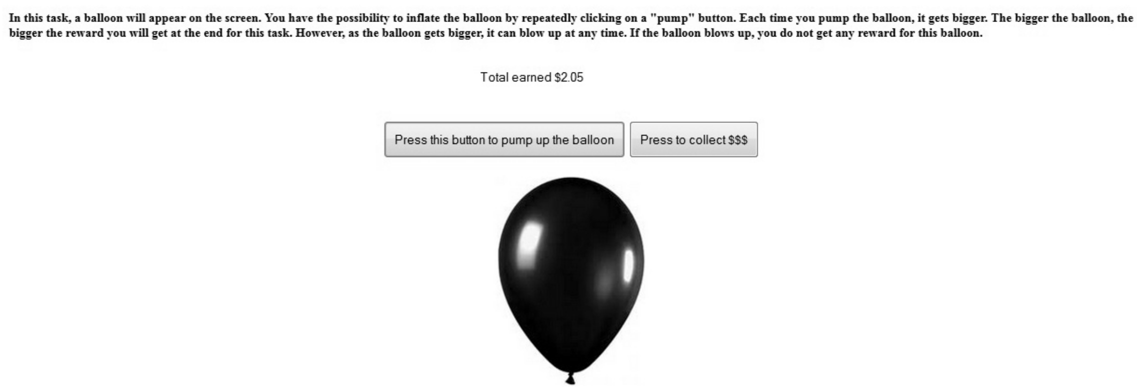
where CE, *certainty equivalence*, is given by the midpoint between the two sure payoffs where the subject switches from the risky option to the sure payoff (6.125 dollars in the examples above), and π denotes the prize in the risky option (7 dollars in our task). If $r > 0.5$, the subject is considered risk averse. If $r < 0.5$, the subject is considered risk loving, and risk-neutral if $r = 0.5$.

We denote a the ambiguity attitude, which is obtained as follows:

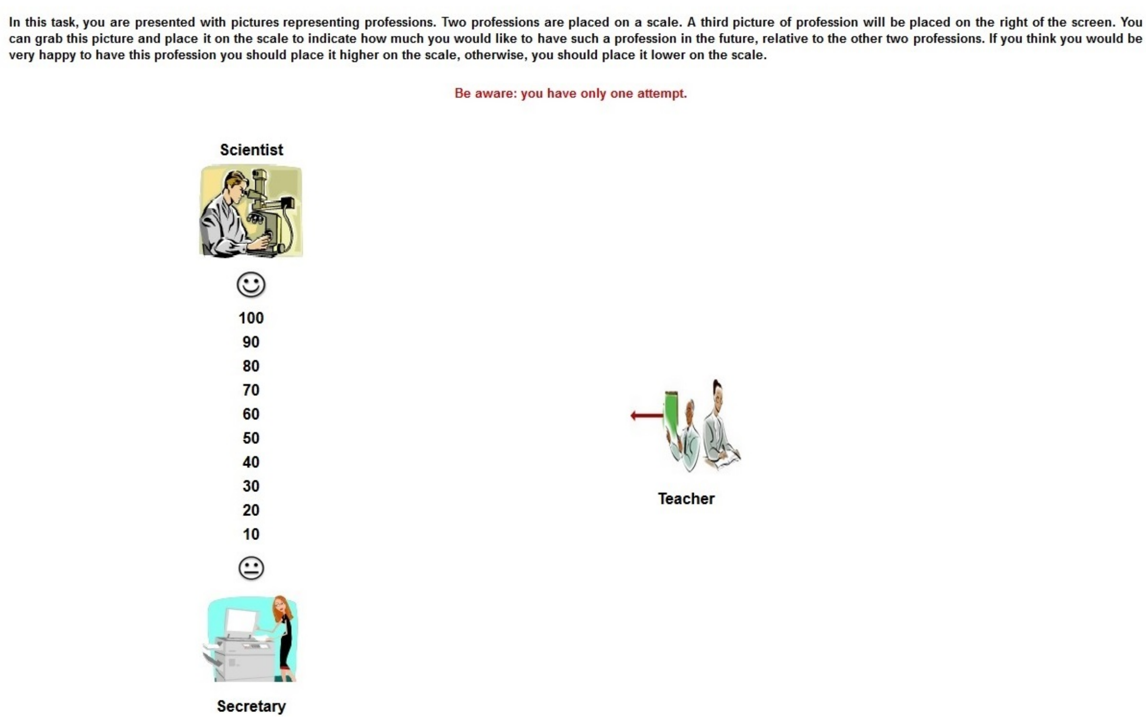
$$a = (CE_r - CE_a)/(CE_r + CE_a)$$

where CE_a is the certainty equivalent in the ambiguous prospect, with $a \in [-1, 1]$. If $a \to -1$, the subject has extreme ambiguity loving. If $a \to 0$, the subject is ambiguity neutral, and if $a \to 1$, the subject has extreme ambiguity aversion.

A.3.3 Task 3: Balloon Analogue Risk Task



A.3.4 Task 4: Aspiration levels



A.4 Regression results on academic scores

Table A.2: IV regression on academic scores

	math	English
Grade 9	-0.113	-0.124
	(0.128)	(0.123)
Relative age	0.039	-0.066
	(0.164)	(0.209)
QLD	-0.022	-0.035
	(0.179)	(0.131)
Constant	0.958	1.174
	(1.183)	(1.119)
Observations	466	468

Observations are clustered by school. Academic scores are standardised to mean 1 and standard deviation 0. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.5 First stage of the IV regression

Table A.3: First stage of the IV regression

	Relative age	
	(1a)	(1b)
Predicted ra	0.498***	0.471***
	(0.043)	(0.041)
QLD	-0.079**	-0.103**
	(0.037)	(0.041)
Grade 9		-0.058
		(0.045)
Male		0.092**
		(0.038)
Math score		-0.014
		(0.018)
English score		-0.003
		(0.016)
Public		-0.084**
		(0.038)
Regional		0.073
		(0.060)
Constant	0.646***	1.205***
	(0.055)	(0.399)
Observations	537	466

Observations are clustered by school. Predicted ra takes value 1 if the participant is born just after the cut-off date and 0 if he or she is born just before the cut-off date. Relative age takes value 1 if the participant is relatively old, and 0 if he is relatively young in his cohort. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B OLS regression results

Table B.1: OLS regressions on performance

	Quasi-experimental sample						Extended sample					
	Piece-rate		Tournament		Difference		Piece-rate		Tournament		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
Relative age	-1.283**		-1.732***		0.447		-0.394	-0.548	-0.611	-0.762**	0.207	0.196
	(0.481)		(0.627)		(0.402)		(0.347)	(0.352)	(0.428)	(0.345)	(0.244)	(0.262)
QLD	-0.626	-0.556	-1.308	-1.552**	0.680*	1.020**	-0.946	-0.658	-1.212*	-1.000**	0.284	0.351
	(0.850)	(0.676)	(0.883)	(0.639)	(0.377)	(0.396)	(0.634)	(0.487)	(0.693)	(0.470)	(0.181)	(0.264)
Grade 9		1.507***		2.040***		-0.557		0.654**		0.837***		-0.213
		(0.519)		(0.424)		(0.395)		(0.302)		(0.290)		(0.222)
Male		0.850*		0.458		0.372		0.356		0.152		0.186
		(0.431)		(0.439)		(0.317)		(0.273)		(0.337)		(0.241)
Constant	7.190***	-5.338	9.530***	-6.578*	-2.336***	1.400	6.999***	1.957	8.348***	2.092	-1.358***	0.135
	(1.623)	(4.389)	(1.671)	(3.629)	(0.713)	(3.146)	(1.211)	(2.982)	(1.308)	(3.084)	(0.352)	(1.919)
N	283	252	283	252	282	251	534	463	532	461	529	458
Add. controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school.
$+ p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Probit regressions for the choice of compensation scheme (stage 3)

	Quasi-experimental sample				Extended sample			
	All		Male	Female	All		Male	Female
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Relative age	-0.078	-0.204	-0.110	-0.282	-0.125	-0.207	-0.102	-0.245
	(0.190)	(0.207)	(0.369)	(0.300)	(0.145)	(0.166)	(0.216)	(0.214)
QLD	0.064	-0.049	-0.129	0.050	-0.136	-0.225	-0.412^{+}	0.077
	(0.212)	(0.219)	(0.343)	(0.336)	(0.119)	(0.150)	(0.229)	(0.230)
Grade 9						-0.041	-0.057	-0.058
						(0.128)	(0.157)	(0.232)
Male		0.403^{**}				0.417^{***}		
		(0.181)				(0.150)		
Tournament performance		0.018	0.023	0.015		0.039^{+}	0.037	0.042
		(0.030)	(0.039)	(0.043)		(0.023)	(0.030)	(0.034)
No. of participants		-0.020^{**}	-0.017	-0.026		-0.017^{***}	-0.011	-0.029^{***}
		(0.010)	(0.013)	(0.016)		(0.006)	(0.008)	(0.008)
Rank guessed at tournament		-0.019^{***}	-0.018^{***}	-0.021^{***}		-0.021^{***}	-0.020^{***}	-0.023^{***}
		(0.004)	(0.005)	(0.006)		(0.003)	(0.005)	(0.006)
N	279	248	126	122	529	455	242	213
Add. controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Average marginal effects; Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Probit for the submission of the piece-rate to a tournament (stage 4)

	Quasi-experimental sample				Extended sample			
	All		Male	Female	All		Male	Female
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Relative age	-0.182	-0.228	-0.447	-0.179	-0.035	-0.104	-0.103	-0.069
	(0.159)	(0.159)	(0.353)	(0.253)	(0.096)	(0.108)	(0.166)	(0.241)
QLD	-0.105	0.067	0.320	-0.168	-0.151	-0.153	-0.012	-0.290
	(0.260)	(0.280)	(0.298)	(0.380)	(0.126)	(0.173)	(0.187)	(0.327)
Grade 9						0.089	0.109	0.146
						(0.119)	(0.172)	(0.181)
Male		0.478**				0.498***		
		(0.199)				(0.148)		
Piece rate performance		0.004	0.051	-0.047		0.016	0.042+	-0.020
		(0.026)	(0.039)	(0.047)		(0.019)	(0.024)	(0.037)
No. of participants		-0.035***	-0.040***	-0.035**		-0.024***	-0.023***	-0.029**
		(0.013)	(0.014)	(0.016)		(0.006)	(0.007)	(0.012)
Rank guessed at piece-rate		-0.021***	-0.025***	-0.017**		-0.019***	-0.020***	-0.018***
		(0.006)	(0.009)	(0.008)		(0.003)	(0.005)	(0.006)
N	281	249	126	123	533	459	241	218
Add. controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Average marginal effects; Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: OLS regressions for the expected rank in the piece-rate (stage 1)

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	7.780^{+}	6.434^{+}	4.874	0.708	-0.683	-1.296
	(3.863)	(3.533)	(4.609)	(2.722)	(2.432)	(3.460)
QLD	0.016	-2.568	-2.783	0.316	-5.284	-5.325
	(2.774)	(4.785)	(4.834)	(2.842)	(3.881)	(3.928)
Grade 9					-6.332***	-6.376***
					(1.986)	(1.984)
Male		-8.285***	-11.497**		-7.045***	-8.079**
		(2.710)	(5.033)		(2.047)	(3.877)
Relative age x Male			4.008			1.334
			(5.365)			(4.363)
Piece rate performance		-2.717***	-2.697***		-2.965***	-2.959***
		(0.433)	(0.427)		(0.313)	(0.310)
No. of participants		0.020	0.029		-0.210	-0.208
		(0.249)	(0.242)		(0.141)	(0.139)
Constant	46.379***	76.530***	77.602***	50.244***	144.424***	145.231***
	(5.186)	(13.999)	(14.646)	(5.223)	(20.660)	(21.169)
N	284	252	252	537	463	463
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area.
Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: OLS regressions for the expected rank in the tournament (stage 2)

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	6.904^{+}	7.083^{**}	4.002	0.575	0.991	-0.828
	(3.533)	(2.975)	(4.296)	(2.549)	(2.201)	(3.244)
QLD	-0.268	-1.042	-1.457	2.714	-0.432	-0.553
	(2.929)	(4.522)	(4.550)	(2.855)	(2.970)	(2.986)
Grade 9					-5.593^{***}	-5.732^{***}
					(1.982)	(1.990)
Male		-12.065^{***}	-18.340^{***}		-7.876^{***}	-10.904^{**}
		(2.919)	(4.994)		(2.268)	(4.032)
Relative age x Male			7.866			3.916
			(5.544)			(4.398)
Tournament performance		-1.481^{**}	-1.459^{**}		-1.803^{***}	-1.792^{***}
		(0.643)	(0.644)		(0.460)	(0.464)
No. of participants		0.282	0.302		-0.023	-0.016
		(0.195)	(0.194)		(0.106)	(0.105)
Constant	47.662^{***}	65.220^{***}	67.385^{***}	45.785^{***}	120.120^{***}	122.599^{***}
	(5.124)	(14.642)	(15.105)	(5.415)	(20.129)	(20.537)
N	284	252	252	537	461	461
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area.
Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: OLS regressions on risk aversion

Relative age	0.068^{+}	0.094^{**}	0.101^{**}	0.035	0.047^{**}	0.045
	(0.038)	(0.045)	(0.048)	(0.024)	(0.023)	(0.037)
QLD	0.025	0.044	0.045	0.021	0.016	0.016
	(0.033)	(0.046)	(0.047)	(0.024)	(0.033)	(0.032)
Grade 9					-0.034	-0.034
					(0.032)	(0.032)
Male		-0.079^{**}	-0.066		-0.076^{***}	-0.079
		(0.036)	(0.072)		(0.025)	(0.050)
Relative age x Male			-0.016			0.004
			(0.068)			(0.057)
N	268	238	238	505	438	438
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: OLS regressions on ambiguity aversion

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	0.012	0.008	-0.028	-0.001	0.005	-0.053
	(0.037)	(0.041)	(0.048)	(0.030)	(0.029)	(0.042)
QLD	-0.010	-0.029	-0.035	0.015	0.013	0.007
	(0.043)	(0.045)	(0.043)	(0.039)	(0.043)	(0.042)
Grade 9					-0.018	-0.022
					(0.019)	(0.019)
Male		-0.003	-0.073		-0.018	-0.110^{**}
		(0.038)	(0.063)		(0.030)	(0.046)
Relative age x Male			0.088			0.121^{**}
			(0.073)			(0.050)
N	264	237	237	502	438	438
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: OLS regressions on the BART

	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	7.763	3.869	5.240	10.276	8.560	14.070
	(14.732)	(14.010)	(15.428)	(9.527)	(10.170)	(13.510)
QLD	0.108	-7.434	-7.217	-0.312	-6.230	-5.713
	(15.120)	(18.623)	(18.710)	(13.786)	(17.565)	(17.786)
Grade 9				1.458	0.944	1.300
				(7.126)	(7.370)	(7.554)
Male		3.563	6.232		8.922	18.034
		(10.180)	(29.130)		(9.691)	(22.719)
Relative age x Male			-3.393			-11.869
			(29.230)			(23.952)
N	259	242	242	488	450	450
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.9: OLS regressions on professional aspirations

	Average rank for the 4 occupations					
	Quasi-experimental sample			Extended sample		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Relative age	-1.463	-2.258	-2.014	0.589	0.128	0.610
	(1.776)	(2.216)	(2.715)	(1.618)	(1.755)	(2.419)
QLD	-1.392	-0.680	-0.643	-2.725**	-2.108	-2.062
	(1.207)	(1.590)	(1.740)	(1.331)	(1.348)	(1.420)
Grade 9				2.065	2.454	2.491+
				(1.621)	(1.480)	(1.475)
Male		-0.189	0.313		-0.170	0.645
		(2.106)	(4.274)		(1.318)	(2.903)
Relative age x Male			-0.629			-1.051
			(5.355)			(3.938)
N	279	250	250	524	461	461
Add. controls	No	Yes	Yes	No	Yes	Yes

Includes control variables for math and English scores, whether the school is public and located in a regional area. Standard errors in parentheses, clustered by school. + $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.1 Instructions script

Thank you for taking part in this experiment. Today you will be engaged in few task. It is fun and you can earn some real money. You will get the money at the end of the session. The decisions you will make in the experiment will determine how much you get at the end. The experiment consists of four tasks. In task 1,2 and 3 you have the possibility to make money. One of these tasks will be randomly selected for payment at the end of the session.

I invite you to listen carefully to the instructions for you to understand the tasks. Please raise your hand at any time if you have a question. Please do not talk to your neighbours during the session.

Task 1

This task will have 5 parts *(write 1, 2, 3, 4, 5 on the board)*.

(Underline 1) In the first part, you will be invite to quickly compute sums of five numbers. Each time you propose an answer (right or wrong) another set of five numbers will appear for you to sum. You will have 3 minutes to solve as many additions as you can. You are not allowed to use a calculator/phone. If this part is selected for payment, you will receive $0.20 per correct answer.

Is it clear?...Is there something which you would like me to explain better?

OK, then please click on the button I understand.

(Underline 2) Let's consider the part 2. In this part the task is exactly the same as in the part 1: you will have 3 minutes to solve as many additions as you can. However, the payment will be different. Now it will depend on whether you are one of the students in this room/school, having the highest number of right answers. If you are in the top third students in terms of number of answers you will receive $0.60 per correct answer. If you are not in the top third of students with the highest number of right answers, you will not get anything. You can see an example of a classroom with 12 participants in the room. In this example only the students ranked 4 or above receive $0.60 per addition correctly solved. Those ranked below 4, do not get anything. In this classroom/school, you are X students, so only those ranked within the top X/3 students will receive $0.60 per correct answer. The others will not receive anything.

Is it clear?...Is there something which you would like me to explain better?

(Underline 3) Let's consider the part 3. In this part the task is once again exactly the same as in the part 1: you will have 3 minutes to solve as many additions as you can. However, you will now have the opportunity to choose how to be paid for it. You can chose either to be paid like in part 1 ($0.20 for each correct answer) or like in part 2 ($0.60 per correct answer only if you are one of the top X/3 students in part 2). You

will be asked to choose between these two possible payment methods before starting to solve the additions.

Is it clear?... Is there something which you would like me to explain better?

(Underline 4) Let's now consider the part 4 of the task. This part is a bit different. In this part you will not have to solve any addition. You will only have to choose a payment method. This payment method will be applied to one of your previous results, the number of right answers that you gave in part 1. This is how it will work: First, you will be informed about how many right answers you gave in part 1. Second, you will have the possibility to choose how to be rewarded for these right answers. You may want to be paid like in part 1, in that case you will receive $0.20 for each correct answer. Alternatively, you will have the opportunity to choose a payment method as in part 2. If you do so, you will receive $0.60 per correct answer if you are one of the top X/3 students in part 1, and nothing otherwise. Importantly this choice has no influence on your payment in part 1 itself *(point on the 1 on the board)*. It will only determine you payment in part 4. Let's consider an example. Imagine you got 5 answers right in part 1, you select competition. Suppose you were not in the top X/3, how much do you get in part 4, if selected? In part 1, if selected?

Is it clear?... Is there something which you would like me to explain better?

(Underline 5) Let's consider the final part of the task. This task is relatively simple. There are no addition to solve or payment method to chose. In this part, you will be asked to guess how well you did relatively to the other participants in the room/school in Part 1 and Part 2. In this classroom/school you are X. This number will be displayed on the screen. You will then be asked to enter a guess about your rank in Parts 1 and 2. You can put any number between 1 (if you think you are the one who has the highest number of correct answers) to X (if you think you are the one with the smallest number of correct answers).

Is it clear?... Is there something which you would like me to explain better? Ok, then please click on the button I understand.

Task 2

The second task consists in two parts *(write 1,2 on the board)*

(Underline 1) Let's consider part 1. You will have to choose whether you prefer safe amounts of money or drawing a ball from Bag A, with a chance of winning $7. Here is the bag. I will now fill the bag with white and orange balls. As you can see there are 6 white balls and 6 orange balls. So you have 50% chance to draw a white ball. (You show conspicuously the balls and place them in the bag). When you decide to draw a ball from Bag A, you draw a ball blindly. If the drawn ball is white, you receive $7.

If the drawn ball is orange, you get nothing. You will be presented with a list of sure amounts of money ranging from $0.35 to $7. For each sure amount you will have to indicate whether you prefer the sure amount, or whether you prefer to draw a ball from the bag. You will have on your screen a decision table which looks exactly like the one you see now on the explanation screen. You will then have to make a decision for each row between drawing a ball from Bag A and a sure amount of money. For example: In the first row, you decide whether you prefer to draw a ball from Bag A and thereby maybe winning $7, or if you prefer taking $0.35 home for sure. In the second row you decide again between drawing a ball from Bag A and a sure amount of money. Now you are offered $0.70 to take home for sure. *(show a printed version of the table and point the different elements when talking)* There is no right or wrong answer. You just have to indicate your preferences. Assume that you very much dislike drawing a ball from Bag A. In this case you might choose the sure amount of money all the time. Assume that you like drawing a ball from bag A very much, then you might check the boxes on the left hand side most of the time. However, even if you like to draw from the bag you may consider not choosing the bag in the bottom row. Do you have an idea why? The easiest way to proceed is for you to make a choice for each row in the table going from the top row to the bottom row. The webpage will help you fill the boxes. Whenever you indicate that you prefer a sure amount to drawing the bag, it will tick all the sure amounts which are higher. If you prefer $2.45 to draw from the bag, you should also prefer $2.8, $3.15 and so on. So the webpage will tick these boxes *(point to them on the print out)*. At the end of the experiment, if this part is selected for payment, only 1 out of the 20 decisions will be used to determine your earnings.

Is it clear?... Is there something which you would like you would like me to explain better? If you do not have any questions, please click on the button I understand.

(Underline 2) Now we explain the second part of the task. The second part is very similar to the first part. The only difference is that Bag A is replaced by Bag B. Now you have to choose between drawing a ball from Bag B and thereby maybe winning $7 or taking a sure amount of money home. Drawing from Bag B works as follows: This bag contains 12 balls. The balls are either white or orange as before, but this time we do not tell you the exact number of white and orange balls. When you decide to draw a ball from Bag B, you draw a ball blindly. If the drawn ball is white you get $7, if t is orange, you get nothing.

Is it clear?... Is there something which you would like me to explain better?

If you do not have any questions, please click on the button I understand.

Task 3

In this task imagine you are blowing up a balloon, just like the one you have on your screen. You have a button labeled 'Press" on your screen. Each time you will click on this button, the balloon will increase in size. The bigger the balloon, the larger the reward will be. Each click on the pump raises your earnings by 5 cents. At each point you can chose to collect the money or to continue pumping. It is your choice to decide how much to pump up the balloon, but be aware that at some point, the balloon will explode. If the balloon explodes before you click on the button Collect $$$, all the money that you accumulated in this balloon will be lost. The balloon can explode at any pump, even at the first one. Also, be aware that all the balloons will not explode at the same time. After you collected the money on a balloon, or after the explosion of a balloon, you will be presented with another balloon. There are 5 balloons in total. The amount accumulated in each balloon will not affect the amount accumulated in the other balloons. If this task is selected for payment, you will receive the earnings accumulated in one randomly chosen balloon.

Is it clear?...Is there something which you would like me to explain better?

Task 4

The last task is different from the precedent ones. Here we will ask you to think about what you would like to do in the future as a professional career. You will see different professions and you will be asked to say how happy you would be to have a given profession relative to two other professions in a comparable area of activity. In the example on your screen, the three professions are Engineer, Factory worker and Technician. Engineer is ranked higher than Factory worker in the classification of occupations and is also associated with a higher social status. The profession of Technician is ranked between the profession of Engineer and the profession of Factory worker in the classification of occupations and is associated with a higher social status. In this task we ask you to think about how happy you would be to have the middle ranked profession (here the profession of Technician), relative to the other two professions. If you think you would be very happy to have this profession you should place it higher on the scale, otherwise, you should place it lower on the scale. In this example, if you would be as happy being a Technician as you would be being a Factory worker, then you should position the Technician close to the Factory worker in the scale. This would indicate that you clearly prefer to be an Engineer over both other professions. If you would be as happy being a Technician as you would be being an Engineer, then you should position the Technician close to the Engineer in the scale. This would indicate that you are indifferent or only weakly prefer the profession of Engineer to the profession of Technician. You will be asked to do this exercise for four different groups of professions.

Is it clear?... Is there something which you would like me to explain better? If you do not have any questions, please click on the button I understand.

You will now start to perform each of the tasks. If you have any question while performing the tasks, please raise your hand and one of us will come to answer your question privately. In each task, you will be given a short explanation on the screen about the task you are about to perform. Please feel free to take the time during the experiment to re-read the instruction of each task before starting it.

B.2 Pictures

[commented out to save space]