WARWICK
ECONOMICS

Small Fish in a Big Pool: The Discouraging Effects of Relative Assessment

Nicholas Bottan & Dan Bernhardt

August 2024 No: 1511

Warwick Economics Research Papers

ISSN 2059-4283 (online)
ISSN 0083-7350 (print)

Small Fish in a Big Pool: The Discouraging Effects of Relative Assessment*

Nicolas Bottan†
Cornell University

Dan Bernhardt‡
University of Illinois

August 9, 2024

Abstract

This paper studies the impact of relative assessment on performance using a quasi-experiment: club-level swimming competitions in the US. By exploiting the age-group structure, where swimmers are assessed against peers within their age group and experience a significant shift in relative standing upon aging up, we identify the causal effects of being assessed against better-performing peers. Using a regression discontinuity design, we find that swimmers, on average, swim significantly slower after aging up. This effect is similar across genders and is most pronounced among swimmers in the middle and top of the ability distribution, while those in the bottom third show no significant change. Our findings highlight the importance of considering the psychological impacts of relative assessment in competitive environments.

Keywords: Relative Assessment, Competition, Performance, Swimming.
JEL Classification: J53, I29, L83, M54, Z2

*We would like to thank John Burbidge, Tom Avischious, and John Pharris from USA Swimming for giving us access to the data. We are also grateful to David Albouy, Robert Bernhardt, Jeff Brown, Nathan Canen, Roch King, Roger Koenker, Corinne Low, Ben Marx, Ed Pacey, Ricardo Perez-Truglia, and Ignacio Sarmiento for their useful comments and discussions. Disclaimer: The views expressed in this paper do not represent those of USA Swimming, the University of Illinois or Cornell University. All remaining errors and omissions are our sole responsibility.

†Brooks School of Public Policy and Department of Economics, Cornell University, New York 14853, United States. Email: nicolas.bottan@cornell.edu.

‡Department of Economics, University of Illinois at Urbana-Champaign, Illinois 61801, United States. Email: danber@illinois.edu.

1 Introduction

How do people respond to being assessed relative to better peers? Do they rise to the occasion and improve performance? Or do they get discouraged and perform worse? Such situations are common in many settings, including sport, education, and the workplace. For example, students admitted to selective high schools or top universities are assessed relative to other high-performing peers who are graded on a curve (Abdulkadiroğlu et al., 2014, Barrow et al., 2020). These concerns also drive the "redshirting" decisions of parents who choose when to enroll children in elementary school (Lubotsky and Kaestner, 2016) or whether to track students by ability (Matthewes, 2021, Figlio and Ozek, 2024). Although such situations are widespread, we lack a good understanding of the effects of relative assessments on performance and how those effects hinge on the individual.

Identifying these effects outside of the laboratory is difficult for many reasons. First, individuals who self-select into more competitive environments likely have different individual attributes and preferences (e.g., Niederle and Vesterlund, 2007, Leuven et al., 2011). Second, even if we overcame this challenge by randomly assigning individuals to more or less competitive environments, they would not only exposed to different levels of competition but also to different qualities of interactions with peers. For example, a student at a top school is assessed relative to higher performing peers, but may also benefit from interacting directly with them (Sacerdote et al., 2011, Zárate, 2023). Moreover, both the learning environment and the curriculum can reflect the ability composition of students (Burgess et al., 2023). A more general problem is that the cost or benefit can hinge on an individual's ability. Conceivably, a particularly strong student may gain from the challenge of being assessed relative to older and smarter peers, while a student who is just "good enough" may suffer (Abdulkadiroğlu et al., 2014). These issues make it difficult to disentangle the effects of relative assessment *per se* from peer effects in most settings.

In this paper, we exploit a quasi-experiment to identify the effect of being *assessed* against better peers on performance, and how those effects vary with an individual's ability. We study the age-group structure of club-level swimming competitions in the United States. This age-group structure allows us to identify the causal effect of changes in the quality of the assessment group against which swimmers are rated. Specifically, we use the institutional features that (i) age groups span *two* years, and swimmers are assessed/rated only within their age group, with prizes in the forms of trophies, medals and ribbons given to better-ranked performers; and (ii) heats for an event (e.g., 100 yard freestyle, short course, for boys aged 11-12) are organized by seed time so that all swimmers in the same heat have similar times, both before and after aging up.

Crucially, unlike in team sports, swimmer age in the United States is measured relative to birthday, not calendar year. The moment a swimmer has a birthday and ages up to a new age-group, they are ranked relative to this new group, regardless of their ability or the timing within the season. Thus, we first see a swimmer who is old relative to their assessment group, and then the swimmer becomes young relative to their assessment group. The most salient aspect of these age group changes is that the assessment group is much faster after aging up. For example, on average, boys aged 11-12 are about 15% faster than boys aged 10 and under—this is true both for swimmers in a given meet and for the performance time standards set by state swimming organizations for championship qualifications and performance standards by USA swimming. Thus, this setting lets us observe an individual both when they are the big fish *and* the small fish within a short time span, and to study their behavior.

We find that, on average, swimmers are discouraged due to being assessed against older and hence faster peers, swimming more slowly; but this discouragement only shows up for swimmers in the middle and top end of the ability distribution. This evidence is consistent with individuals slowing due to psychological discouragement rather than the outcome of a "rational" effort-cost-benefit calculation.

We use data from the universe of club swimming competition times in the United States to follow individual swimmers over time and obtain their exact date of birth. We focus on a sample of swimmers who compete within a window of days both before and after a birthday when they would change age group, ("*aging up*"). We evaluate the immediate effect of aging up on performance using a regression discontinuity design, where we include swimmer-event fixed effects to control for event-specific ability. We focus on swimmers ages 10 to 14, the prime ages of club swimming, where the shock of aging up is the largest.[1]

This context has several features that let us address the shortcomings of studying the effects of relative assessment outside the laboratory. First, there is a large discrete change in the quality of the assessment groups after aging up. Swimmers cannot sort into easier or harder categories. Swimming events are standardized, and we can compare performance between swimmers and events. Second, there are very small differences in what can be thought of as the effects of *direct* interactions with peers. This is by design due to how meets are arranged: swimmers in an event are sorted into heats by seed time. As a result, *within* a heat, swimmers are usually of very similar ability. Thus, a swimmer who ages up is shifted down in the overall distribution—there are just typically *more* swimmers who are faster, i.e., more of the other heats have faster swimmers. Third, relatedly, from the perspective of other swimmers in the new (older) age group, adding one more (younger) swimmer to a heat also does not affect them. Fourth, because we see a swimmer's times history and have a large sample size, we can construct measures of ex-ante ability to study heterogeneous responses according to ability.

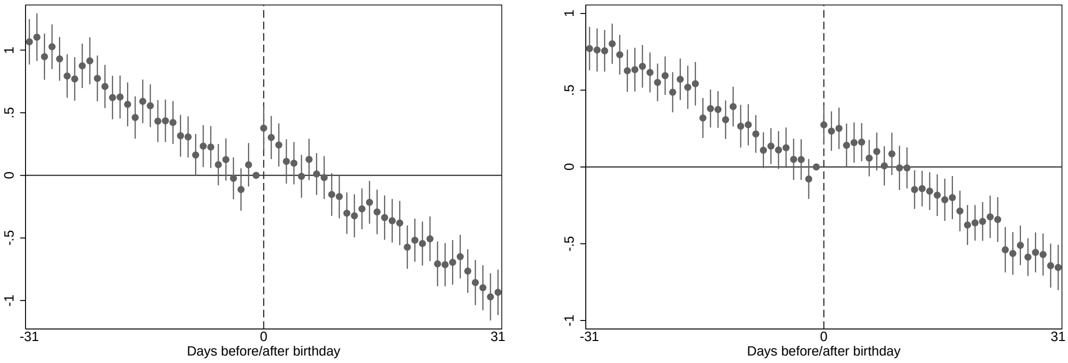
We first document that there are indeed large changes in the quality of competition after aging up. On average, peers are 15% faster (over a standard deviation faster) when turning age 11 and 7.5% faster when turning 13. These changes represent a shift in the overall distribution rather than just a movement in one of the tails: competition is significantly faster in both the upper and lower tails. Next, we examine the effects of being assessed against better peers on performance, measured by their recorded times at meets.

Figure 1 conveys the essence of our findings, showing that swimmers slow down sharply after aging up. On average, they swim around 0.3 to 0.4 percent slower (about 14% of a within swimmer-event standard deviation). The figure reveals that these effects are long-lived within the month window examined, represented by an upward, roughly parallel, shift after aging up. Similar patterns hold over 90 day windows. We find similar discouragement effects for boys and girls, consistent with laboratory experiments suggesting that men and women respond similarly to competition restricted to the same gender

[1]Also, after age 14 many swimmers leave club swimming for high school swimming—not included in our data.

Figure 1. Swim time relative to birthday

(a) Age 11 (b) Age 13



Notes: Sample of swimmers competing in meets within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. First day of meet used as reference time. Dots are point estimates of running variable dummies (-1 is the omitted category) from an OLS regression that includes swimmer-event fixed effects. Dependent variable is the logarithm of swim time (in seconds) times 100. Controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. Ticks indicate 95% confidence intervals, standard errors are clustered at swimmer level.

(e.g., Niederle and Vesterlund, 2007).

We rule out the possibility that the results are driven by other unobserved factors related to a birthday rather than by changes in competition. We replicate our analysis using non-group-change birthday cutoffs (e.g., 10, 12 and 14), meaning that assessment groups are held constant during these birthdays. We find precisely estimated zeros, ruling out the possibility that our results are spuriously driven by unobservable changes in behavior at a birthday.

Two main mechanisms can lead swimmers to slow down when assessed against better peers. First, it could be a rational response from trading off the marginal benefits against the marginal costs of exerting effort. For instance, classical tournament theory would predict such a response if the likelihood of winning became small (Connelly et al., 2014). We find evidence that is inconsistent with this mechanism. First, swimmers slow down by roughly the same amounts in their best strokes—where their placement chances are most affected—as in their other strokes. Second, high-ability swimmers—whose placements chances are most affected by aging up—slow down by the same amounts as middle-ability swimmers. That is, differences in the changes in stakes do not affect the degree to which swimmers slow down. Only swimmers in the bottom third or so of the ability distribution are not affected by aging up.

Our preferred interpretation is that swimmers get discouraged. The discouragement effect could be driven by different factors. For instance, self-image concerns can negatively affect performance. This could be due to changes in perceived self-efficacy (Bandura, 1977) or to lower self-esteem resulting from more negative comparisons with the much faster competition (Kuhnen and Tymula, 2012). These theories are consistent with our finding that the slowdown is similar in magnitude throughout the ability distribution, except for the unaffected lower tail of swimmers who were not competitive prior to aging up.

Our findings are also consistent with individuals trying to maximize their "ego utility" and thus changing their effort choices to manage their self-image (Köszegi, 2006, Bénabou and Tirole, 2002). As swimmers age up into a new group, their reference points change making them perceive their performance as worse (i.e., they see themselves as 'losing' relative to their new peers) leading them to reduce their efforts to protect their self-image. Our findings are also consistent with theories of aspirations suggesting that having too high of aspirations relative to one's current standard may lead to frustration and lower investment (Genicot and Ray, 2017).

Our paper relates to a broader literature documenting the importance of positional externalities on behavior. Both observational and experimental work has documented the importance of relative concerns on a variety of behaviors and contexts such as laboratory settings (Kuziemko et al., 2014), labor supply and performance (Bandiera et al., 2005, Kuhnen and Tymula, 2012, Barankay et al., 2012, Rosaz et al., 2016), people's location choices (Bottan and Perez-Truglia, 2022), premium credit card take-up (Bursztyn et al., 2018), and risk-taking among wartime pilots (Ager et al., 2022). We complement this work by using field data, highlighting how a large negative shock to ones relative assessment in competitive settings can lead to discouragement and reduced performance.

Our paper identifies a potential mechanism that shapes educational outcomes at various stages of schooling. There is extensive research showing that students who are relatively younger than their peers when starting school have worse outcomes and these effects are long-lasting (e.g., Bedard and Dhuey, 2006, Elder and Lubotsky, 2009, Lubotsky and

Kaestner, 2016, Matta et al., 2016, Black et al., 2011). Relatedly, work documenting the "elite illusion" (Abdulkadiroğlu et al., 2014, de Roux and Riehl, 2022, Barrow et al., 2020, Matta et al., 2016, Ribas et al., 2020) finds that marginally admitted students to selective institutions receive lower grades, are less likely to attend college, and suffer negative impacts on socio-emotional traits (Fabregas, 2023, Dasgupta et al., 2022). At the college level, studies that compare outcomes between students admitted to the *same* set of colleges reach a counter-intuitive conclusion: university quality does not matter for future earnings (Dale and Krueger, 2002, 2014, Mountjoy and Hickman, 2021). These different contexts share a common thread: the underperforming students are the small fish in a large pond. Our paper suggests the possibility that large discouragement effects from relative assessments can account for these findings.[2]

This paper also contributes to a literature on the psychological and behavioral drivers of performance in sports. Genakos and Pagliero (2012) document that individuals just behind in weightlifting championships take more risks and perform better, while interim leaders perform worse. Berger and Pope (2011) find that teams slightly behind at halftime in NBA games are actually more likely to win. Brown (2011) shows that golfers perform worse in the presence of a superstar at the tournament, while Pope and Schweitzer (2011) documents evidence of loss aversion in putting behavior. There is also evidence on the importance of reference points in baseball and marathon running (Pope and Simonsohn, 2011, Allen et al., 2017). Researchers also document psychological "choking" under pressure e.g., in tennis on serves (Paserman, 2023), archery (Bucciol and Castagnetti, 2020) and free throw shooting in basketball (Toma, 2017). Our results add to the growing evidence on the importance of psychological mechanisms, as we show that relative assessments can generate adverse effects not captured by classical rational models.

[2] For example, Mountjoy and Hickman (2021) find that students who gain admission to both University of Texas-Austin and UT-Permian Basin earn the same regardless of the school attended. This selected sample of students is clearly at the top of the UT-Permian Basin's cohort but far from the top at UT-Austin. Our analysis suggests that being assessed relative to better peers at UT-Austin could discourage those students, causing them to perform less well and learn less, reducing their earnings to levels similar to what they would have obtained had they attended UT-Permian Basin.

2 Background: Swimming in the U.S.

2.1 Institutional

Swimming competitions in the United States are governed by USA Swimming. The organization has a membership of over 400,000 swimmers that includes all levels of competitors, ranging from age-group to Olympic swimmers.[3] To compete, a swimmer must join one of the over 2,800 clubs around the country. Training is organized independently by clubs and is typically arranged by swimmer ability.

USA Swimming sanctions many types of competitions. By far the most common type is club Invitational Meets, where a host club invites several other clubs to participate. Such competitions account for almost 97% of swim times registered with USA Swimming. Most Invitationals do not have entry time standards, allowing all club members to participate. At these meets, swimmers may register times that qualify them for (e.g., state) championship competitions. At most meets, swimmers participate in multiple events. Events are defined by the combination of distance, style, gender, age-group and course type (e.g., 100 Breaststroke, male, 11-12, Short Course Yards). There are two broadly-defined seasons where different courses are used. During the winter, most competitions take place indoors in pools with lengths that follow the Short Course Yards standard (a 25 yard length). During the summer, most competitions are over the Olympic length of 50 meters (Long Course Meters). Most invitational meets are three-day events.

2.2 Age-group swimming

Before competing in a meet, coaches submit their teams' registrations, which contain each swimmer's seed times several weeks in advance (the timing varies depending on the meet and when coaches enter the data). Seed times are designed to approximate how quickly a swimmer will swim in a race for comparable ranking purposes. A seed time typically corresponds to a swimmer's best previous time in an event. Along with the seed time, coaches submit a swimmer's age measured at the first day of the meet. Because

[3] High school team swimming competitions are not part of USA Swimming.

a swimmer's date of birth is part of their USA Swimming id number, coaches cannot misreport ages. Races are organized by event (e.g., 100 Freestyle, boys) and age group.

One needs to distinguish between how races are scored and how heats are organized because they do not always use the same age cutoffs. Swimmers are almost always scored by the standard age group cutoffs (10 and under, 11-12, 13-14), but are sometimes grouped differently in the races themselves, particularly ages 13-16.[4] The top swimmers in each age group/event are usually awarded ribbons or medals.

The race-age structure used in each meet is not recorded in our data. To understand this, we selected a random sample and manually inspected the information sheets for a sample of non-championship meets from the 2015-2016 season. In total, we coded 224 meets, taking note of the age groupings used to form heats. Table A2 shows that about 70% of meets group 11-12 year olds, using the standard cutoffs for that age group. In contrast, only 36% of meets for 13-14 year olds are grouped this way. Instead, most meets group this age group with older swimmers forming 13 and older heats (almost 45%) and only 32% of meets use the age 15 cutoff. As a result our analysis focuses on ages 10-14.

Lane assignments are based on the ranking of seed times within a race-age grouping. For example, with an 8 lane pool and 76 swimmers in a grouping, swimmers ranked 1-8 by time are commonly grouped together in a heat, swimmers 9-16 are grouped, and so on, with a rump heat of 4 swimmers consisting of those who either do not have a seed time or who are the slowest.[5] Thus, a swimmer ranked 40^{th} would be the slowest in their heat, while a swimmer ranked 41^{st} would be the fastest in their heat. As a result, a swimmer's ranking within a heat is roughly uniformly distributed. Within a heat, lanes are assigned from the center moving outwards,[6] and since ranking within a heat is roughly uniformly distributed so is lane assignment.

This structure has implications for our research design and interpretations. *From a*

[4]Championship meets are organized strictly by standard age groups. In Invitational meets, ages 13-16 are sometimes scored together, in which case the improvement in competition upon turning 13 is even greater than that associated with comparisons to 13-14 year olds, implying that our measure of the degree of improvement is conservative.
[5]With "circle seeding" lane assignment of the fastest heats is slightly different, but qualitatively similar.
[6]For example, in an 8 lane pool, lane 4 is assigned to the number 1 seed, lane 5 to number 2, then lanes 3, 6, 2, 7, 1, 8.

ranking perspective, swimmers who turn 11 or 13 immediately face assessment against older swimmers who have more experience and physical advantages. In sharp contrast, *within a heat*, there are typically very small differences in the swim times of different swimmers, and the randomness in the number of swimmers in an event means that most swimmers are equally likely to be in a heat with slightly faster swimmers as in a heat with slightly slower swimmers. Together, this means that swimmers who age up may face large falls in how they rank within an age group, but aging up has minimal systematic effects on the extent of competition that a swimmer faces within their heat.

2.2.1 Effects of Relative Assessment

Exploiting exogenous variation in the composition of competition is not unfamiliar in education research. Many papers exploit different sources of plausibly exogenous variation in class composition to examine its effects on test scores (see Sacerdote et al. (2011) for a survey). A shortfall of this research design is that changes in class composition also presumably affect direct peer interactions, which, in turn, would affect the estimated effect of class composition.

Our setting does not face this problem due to how races are organized. Swimming features two types of direct peer interactions. First, there is the effect of coaches and swimmers training under a given coach. Interviews with coaches reveal that changing coaches is largely unrelated to aging up, but rather is related to a swimmer's performance (i.e., swimmers are grouped for training based on ability). The second source of direct peer interaction is the swimmers assigned to lanes directly adjacent to their own lane. From the heat and lane assignments, these direct competitors are close in ability and are unlikely to change significantly or vary systematically upon aging up.

2.3 Time measurement

Almost all times are measured electronically using pressure pads at the ends of the pool that activate when touched by a swimmer. Timekeepers serve as backups in case of a faulty read or touch. All times are collected and submitted digitally to USA Swimming to be processed and included in the SWIMS time database. USA swimming maintains and constantly updates the times database. Times are publicly available through their website, where one can search by swimmer name and view all of their recorded times (e.g., one can see the race times for Michael (Fred) Phelps). Even though extensive detailed information is collected at meets (including lane assignments, heat number, seed time, prelim or final), most of this information is lost when processed for the SWIMS database.[7]

We have access to the SWIMS times database, containing all registered swim times from competitions until June 2015 (over 80 million times). Although the data was anonymized, swimmers can be followed over time by their unique ids. For each swimmer, we have data on their gender, date of birth and club (if affiliated). In addition, for each swim time we know the swim date, Meet (and meet type), Event (e.g., 100 Freestyle, Long Course Meters), Time (measured with a precision of a hundredth of a second), Standard, and Hi-Tek Points (a score designed to be comparable across ages, events and gender).

Swimmers start competing at a young age—55% start by age 10—and 58% are girls. Most times (72%) are recorded between the ages of 10 and 15, and swimmers compete the most at age 12. On average, swimmers compete in 10 meets a year, although there is significant heterogeneity: the 25th and 75th percentiles are 6 and 14 meets a year. On average, a swimmer records 5 times at a meet. The database does not include heat or disqualified swimmers (e.g., for an early start), or information on whether a time corresponds to a prelim or final time for meets that use this structure.

3 Identification Strategy

To identify the causal effect of relative assessment on performance, we exploit the discontinuous changes in age group at the relevant birthdays. Because we observe the universe of swim times, we can focus on a sample of swimmer-events around the time of their birthday. In particular, swimmers change age groups just after their 11th and 13th birthdays.[8]

[7]Appendix Figure A.1 presents a screen-shot of the online time search results.

[8]A swimmer who "ages up" mid-way through a multi-day meet remains the same (younger age) throughout the meet, so age is assigned relative to the first day of a meet, i.e., running time t relative to a birthday is assigned to be -1 for a swimmer who ages up midway through a meet.

Because we can track swimmers just before and after aging-up, the setting naturally lends itself to a regression discontinuity design, where the running variable is the first day of a meet in relation to their birthday.

For a given outcome $Y_{i,e,t}$ (e.g., the logarithm of race time times 100), for individual i competing in event e on day t (relative to their birthday at $t = 0$) a naïve model would be:

$$Y_{i,e,t} = \beta \mathbb{1}(t \geq 0) + f(t) + \mathbb{1}(t \geq 0) * f(t) + \phi + \varepsilon_{i,e,t}, \tag{1}$$

where $f(t)$ is a function that is free to differ on either side of the birthday and β estimates the local average treatment effect of changing age group. The naïve model would include event-specific fixed effects (ϕ) to account for different levels in the outcome for different events, and the error term, $\varepsilon_{i,e,t}$, would be assumed to be uncorrelated with the variables of interest, which, in this naïve model, would have to be assumed independent (inappropriate given our panel of swimmer-events).

This naïve model has problems: (1) it ignores the panel structure of the data and assumes that observations are independent; (2) the estimate of β would likely be biased if swimmers behave strategically around the threshold (e.g., quit right after aging up); (3) it implausibly assumes that the evolution of times is the same across different events; (4) it does not account for other potential changes that may happen due to moving into a more competitive category (e.g., swim less events).

To address these potential problems, we focus on a sample of swimmers who compete throughout a window around their birthday, and apply restrictions that minimize sample selection biases (described in detail in Section 3.1). We estimate the following model:

$$Y_{i,e,t} = \beta \mathbb{1}(t \geq 0) + f_e(t) + \mathbb{1}(t \geq 0) * f_e(t) + X_{i,t} + \eta_{i,e} + \varepsilon_{i,e,t}, \tag{2}$$

where we include *swimmer-event* level fixed effects ($\eta_{i,e}$) to estimate changes within both swimmer-events. For example, a swimmer might be better in breaststroke than backstroke, or in longer distances than in shorter ones. This lets us fully exploit the panel structure of the data and address individual-event level selection. We allow the function of the time variable to differ not only on each side of the discontinuity, but also by event ($f_e(t)$). To account for other potential changes that may affect swimmer performance, we include a vector of swimmer-meet characteristics ($X_{i,t}$) that includes the number of heats and distance swum during the day, and the number of days since the last race. Standard errors are always clustered at the swimmer level.

3.1 Sample construction

To estimate our baseline model, we first define a 31 day window both before and after a swimmer's birthday. This window width ensures that we capture enough swimmer-event times on each side of the threshold. We obtain similar results using larger window sizes.

We require at least one swimmer-event observation on each side of the threshold (i.e., we observe the swimmer in a particular event both before and after their birthday, when they compete in the next age group). This restriction minimizes selection biases that could be induced by changes in the composition of swimmers or by changes in the composition of events. For example, if the worst swimmers stop competing right after aging up in order to avoid facing the harder assessment group, then even if the true effect were zero, the change in composition could bias our estimated coefficient downward. We address this by guaranteeing that the same swimmer-events are observed both before and after a swimmer ages up and by including swimmer-event fixed effects in regressions. This restriction excludes swimmers who may have dropped out (either permanently or temporarily). To the extent that this sub-sample drops out due to anticipating the potential negative effects of facing stronger competition, our results represent lower bounds on the (negative) effects of relative assessments. We also restrict our sample of swimmers to those who have been competing for at least one year to better measure swimmer ability.

Appendix Table A3 presents descriptive statistics for our age 11 and 13 birthday samples. By applying tighter restrictions, the sample changes in the expected direction: the remaining swimmers compete in more meets, swim more often, began swimming at a younger age, and are better swimmers (e.g., measured by 100 Freestyle time). Of note, the average club quality (measured by a club's average Hi-Tek score) does not change significantly. When looking at average swim times across events, one must be careful when comparing across columns because the composition of events may change (less experienced swimmers tend to compete in shorter distance events). We drop events that are not swum consistently over time (mostly very long distance races such as the 1650 yard race).

4 Results

4.1 Quantifying the change in competition after aging up

We first document the extent to which the quality of competition improves upon aging up. To do this, we calculate the average swim time of swimmers' competitors in the events they swim using the standard age group definitions.[9] Figure 2 presents graphical evidence of the evolution of the logarithm of average competitor times (times 100), controlling for swimmer-event fixed effects and day-of-race varying controls. We see sharp improvements in the competition—average times of competitors drop by about 15% and 7.5% when turning 11 and 13, respectively. Moreover, our measure at age 13 represents a lower bound on the improvement because many meets define the category as 13-16 rather than 13-14.[10]

Table 1 presents estimates of equation (2) for different (linear, cubic) functional forms of time, and presence or absence of swimmer-meet controls. The quality of competition improves substantially upon aging up, reflecting that the older swimmers in the new age group are more experienced and have sizable physical advantages. When moving into the 11-12 age group, the competition is about 15% (s.e. 0.0916) faster on average. At the age 13 cutoff, the effect is smaller, with the competition being 7.5% faster (s.e. 0.001).

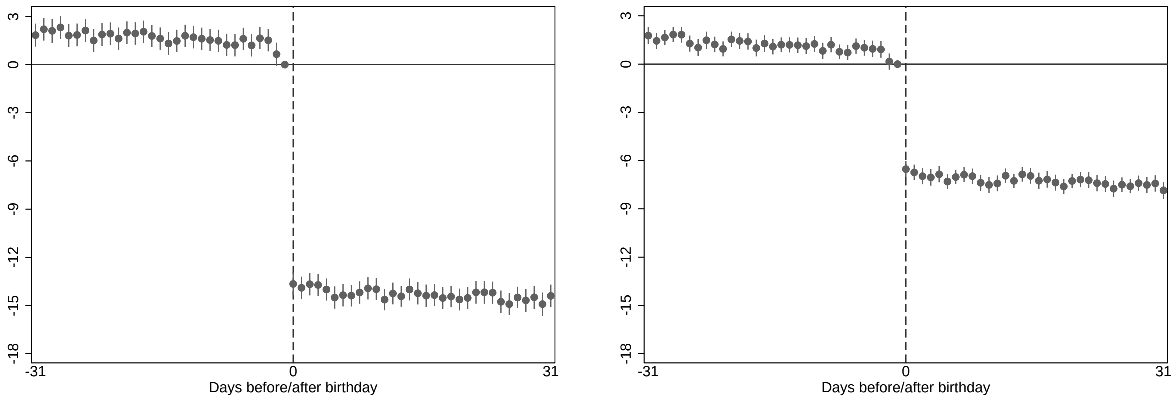
The changes in average competition represent a shift in the distribution, not just a change in the tails. We repeat the same analysis, but examining effects at different moments of the distribution: the 5th and 95th percentile in columns (5) and (6). While both tails of the distribution shift, the change is larger at the bottom of the distribution (i.e.,

[9]Specifically, for the 11 cutoff we group 10 and under when swimmers are 10, and 11-12 when they age up. For the age 13 cutoff, when a swimmer is 12 we calculate the 11-12 average, and when the swimmer is 13 we calculate the 13-14 average.

[10]See Appendix Figure A.2 for a figure that uses the 13-16 definition instead. On average, times of competitors are about 9.5% lower using the 13-16 definition.

Figure 2. Average competition time relative to birthday

(a) Age 11 (b) Age 13



Notes: Sample of swimmers competing in meets within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. First day of meet used as reference time. Dots are point estimates of running variable dummies (-1 is the omitted category) from an OLS regression that includes swimmer-event fixed effects. Dependent variable is the logarithm of competitor's times excluding own time (in seconds) times 100. Controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. Ticks indicate 95% confidence intervals, standard errors are clustered at swimmer level.

Table 1. Changes in Average Competitor's time when changing age-group

	(1)	(2)	(3)	(4)	(5)	(6)
					Competition Percentile	
Dep. Var.:	log(Average Competitor's Time)*100				5th (slow)	95th (fast)
	Panel A: Age 11					
Treatment	-15.27***	-14.76***	-15.18***	-14.48***	-17.94***	-10.58***
	(0.107)	(0.247)	(0.102)	(0.229)	(0.175)	(0.0706)
Obs	441,915	441,915	441,915	441,915	441,915	441,915
Nr Swimmers	39,750	39,750	39,750	39,750	39,750	39,750
	Panel B: Age 13					
Treatment	-7.651***	-7.622***	-7.507***	-7.402***	-9.319***	-5.173***
	(0.0829)	(0.185)	(0.0807)	(0.175)	(0.146)	(0.0564)
Obs	450,291	450,291	450,290	450,290	450,290	450,290
Nr Swimmers	45,439	45,439	45,439	45,439	45,439	45,439
Controls	No	No	Yes	Yes	Yes	Yes
f()	Linear	Cubic	Linear	Cubic	Linear	Linear

Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of changing age-group. The functions of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.

among the slowest swimmers). For instance, at age 11 the 5th percentile (slowest end of the distribution) is almost 18% faster, while the 95th percentile (fastest) is 10.6% faster.

4.2 Effects on own performance

We next examine how aging up affects a swimmer's performance. Figure 1 presented graphical evidence showing the evolution of the adjusted logarithm of own swim time (multiplied by 100) around the 11th and 13th birthday. On average, a swimmers' performance improves roughly linearly over time. However, we observe sharp discontinuous increases in their times (i.e., they swim slower) exactly on their birthday, when transitioning into an older age group. This increase corresponds to an upward shift in the trend (and hence a parallel shift in times) rather than a temporary jump followed by a return to the previous improvement trend.

We quantify the change in performance in Table 2, where we present estimates of equation (2). At both the 11^{th} and 13^{th} birthdays (panels A and B) a swimmer's own time increases after aging up. This effect is highly statistically significant (p-value$<$0.001), even after conservatively using a Bonferroni correction.[11] The estimates in column (1) imply that swimmers are 0.340% slower (s.e. 0.0271) when turning 11 and 0.334% slower (s.e. 0.0223) when turning 13. Estimates are similar in significance and magnitude when we use alternative functional forms for time (columns (2) and (4)) or add additional controls (column (3)).

Although the point estimates seem small, the effects are economically/strategically meaningful: the change represents an increase of almost 14% of a standard deviation within swimmer-event at age 11 and 13.4% at age 13. To provide an alternative benchmark, we consider the 100 yard freestyle event (short course), the most popular event swum, where swimmers slow down on average by 0.487% (s.e. 0.138) when turning 11, and by 0.407% (s.e. 0.106) upon turning 13.[12] In competitions with at least 20 swimmers, the average slow down by 11 year olds exceeds the margin between third and fourth 30% of the time; and the average slow down by 13 year olds exceeds the margin between third

[11] We adjust by multiplying p-values by six to account for the six different birthdays (10–15) tested.
[12] See Appendix Table A5 for effects by event.

Table 2. Effects of age-category change on performance

	(1)	(2)	(3)	(4)
Dep. Var.: log(Swim Time)*100				
	Panel A: Age 11			
Treatment	0.340***	0.427***	0.356***	0.465***
	(0.0271)	(0.0608)	(0.0271)	(0.0604)
Obs	441,920	441,920	441,920	441,920
Nr Swimmers	39,750	39,750	39,750	39,750
	Panel B: Age 13			
Treatment	0.334***	0.332***	0.329***	0.326***
	(0.0223)	(0.0475)	(0.0228)	(0.0472)
Obs	450,295	450,295	450,294	450,294
Nr Swimmers	45,439	45,439	45,439	45,439
Controls	No	No	Yes	Yes
f()	Linear	Cubic	Linear	Cubic

Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of changing age-group. The functions of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.

and fourth 32% of the time, i.e., the difference between medaling and not.[13]

Table 3 shows that our main finding that swimmers slow down immediately after changing age group is robust to many checks. Estimates are similar when using wider windows of 84 or 42 days before and after a birthday. One concern could be that when aging-up, swimmers change the composition of events swum. Although we avoid this issue by construction—by focusing on swimmer-events consistently swum on both sides of a birthday—one could imagine that they start swimming new longer events that could generate spill-overs to other races. This is implausible for two reasons. First, our main analysis already controls for meet and day of race controls that include number of races and distance swum. Second, in column (4) of Table 3 we drop all swimmers who ever

[13] See Appendix Figure A.3 for the CDFs of third to fourth place gaps in races with more than 20 swimmers.

compete in a longer distance race (defined as greater than 800 yards or meters) within this window. The estimated effects remain similar in significance and magnitude, indicating that this does not drive the increases in times after a birthday.

Table 3. Effects of age-category change robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
*Dep. Var.: log(Swim Time)*100*								
	Panel A: Age 11							
Treatment	0.319***	0.338***	0.356***	0.339***	0.361***	0.311***	0.288***	0.395***
	(0.0160)	(0.0183)	(0.0271)	(0.0281)	(0.0281)	(0.0406)	(0.0366)	(0.0318)
Obs	744,119	616,871	441,920	415,170	411,707	330,817	191,686	312,678
Nr Swimmers	39,750	39,750	39,750	38,046	37,712	33,460	20,610	34,489
	Panel B: Age 13							
Treatment	0.274***	0.314***	0.329***	0.281***	0.336***	0.329***	0.289***	0.366***
	(0.0135)	(0.0153)	(0.0228)	(0.0276)	(0.0237)	(0.0340)	(0.0309)	(0.0270)
Obs	749,594	626,216	450,294	315,543	412,521	343,730	197,865	289,320
Nr Swimmers	45,439	45,439	45,439	34,895	42,754	38,502	23,583	37,074
Window Size	± 84	± 42	± 31	± 31	± 31	± 31	± 31	± 31
Other	–	–	–	No long dist.	Regular	Donut	Multi Event	Large meets

Notes: Sample of swimmers competing within given days (both before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. All estimates include controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of changing age-group. All estimates use a linear function of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Window Size* indicates the sample restriction for the number of days before and after a birthday swimmers must compete within. *Other* indicates other sample restrictions. In column (4), *No long dist* drops swimmers who ever participate in events longer than 800 yards or meters. Column (5), *Regular* drops major championship meets. Column (6), *Donut hole* drops observations within 4 days before and after cutoff. Column (7) restricts to swimmers participating in more than 2 meet-events before and after aging up. Column (8) restricts to races with more than 20 participants. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.

We also address concerns that our results may reflect strategic tapering for a major meet (e.g., a state championship). Swimmers taper by reducing the intensity of training prior to major competitions. Thus, a concern is that the club tapers prior to a major competition with time standards that a swimmer only meets prior to aging up (so that a taper just after a birthday becomes irrelevant). These major competitions right before a birthday could conceivably drive the jump we observe around a birthday. If this were so, our effects should disappear once we drop major competitions from our sample and only focus on "regular" meets. Column (5) provides evidence against tapering driving our effects—in fact, the coefficient is slightly *larger* (0.361% at age 11) and remains highly statistically significant (s.e. 0.0281). Column (6) shows that findings are qualitatively unchanged when we estimate the model using a 4-day donut hole around the birthday. Column (7) restricts the sample to swimmers-events that appear more than once on either side of the birthday.

Finally, column (8) restricts the sample to large meets, defined as there being more than 20 swimmers in a given event. This is important as smaller meets are less competitive and could attenuate the estimates of the effects of changing age group. Indeed, the estimated coefficients are over 10% larger than the baseline estimates in column (3). Taken together, this suggests that the effects we find are driven by changes in the extent of ranking competition rather than other potential factors related to a swimmer's birthday or changing age group (e.g., swimming longer events).

An alternative way to evaluate the effect of improvements in competition on performance is by directly examining the relationship between average competition and own performance. Examining this relationship is difficult because better swimmers can and sometimes do select into better meets. To overcome this identification challenge, we exploit the variation created by changes in age groups and use it as an instrument for the average quality of the competition. This way, assuming that all of the effect of aging up operates through the quality of competition, we can directly quantify the effect of changing competition on performance. We present results in Table 4, where the dependent variable is the logarithm of a swimmer's time and the endogenous variable is the logarithm of the average competition time. Consistent with our reduced form findings, swimmers swim slower when assessed against faster competition. For example, if the average time of the competition improves by 1%, a swimmer performs almost 0.04% worse (s.e. 0.003) at age 13 in Panel B, column (1) (i.e., a behavioral elasticity of 0.04).

It is useful to contrast the magnitude of these average effects with those of other studies. For example, in the context of peer quality in education, our results align with lower bound estimates from the literature as in Arcidiacono et al. (2012), Burke and Sass (2013)) who find that a one standard deviation in peer quality increases math scores by 3% (4%) of a standard deviation. Similarly, Cornelissen et al. (2017) find that a one standard deviation increase in peer ability in the workplace increases *wages* by 0.3% on average. Our instrumental variable estimates suggest effects of a similar magnitude on average: making the average competition one standard deviation faster (8.78% at age 13) worsens a swimmer's performance by 0.349%. It is important to highlight that, if anything, our estimates should be interpreted as lower bounds as we are only identifying the *immediate* effect of relative assessment (and we drop swimmers who are so discouraged that they do not swim in the month after aging up).

Table 4. 2SLS Estimates of Change in Competitor's Time on Own Performance

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.: log(Swim Time)*100						
	Panel A: Age 11					
log(Avg. Comp.)*100	-0.0234***	-0.0321***				
	(0.00181)	(0.00426)				
log(5th ptile Comp.)*100			-0.0198***	-0.0274***		
			(0.00155)	(0.00368)		
log(95th ptile Comp.)*100					-0.0336***	-0.0470***
					(0.00259)	(0.00622)
Obs	441,915	441,915	441,915	441,915	441,915	441,915
Nr Swimmers	39,750	39,750	39,750	39,750	39,750	39,750
	Panel B: Age 13					
log(Avg. Comp.)*100	-0.0433***	-0.0461***				
	(0.00312)	(0.00696)				
log(5th ptile Comp.)*100			-0.0353***	-0.0379***		
			(0.00259)	(0.00586)		
log(95th ptile Comp.)*100					-0.0635***	-0.0679***
					(0.00455)	(0.0102)
Obs	450,290	450,290	450,290	450,290	450,290	450,290
Nr Swimmers	45,439	45,439	45,439	45,439	45,439	45,439
f()	Linear	Cubic	Linear	Cubic	Linear	Cubic

Notes: Sample of swimmers competing within 31 days (both before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate 2SLS regression where the instrumented endogenous variable, the logarithm of the average competitor's race time times 100 (or the 5th or 95th percentiles), is instrumented by the treatment indicator for change in age-group. All estimates include swimmer-event fixed effects and controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. The functions of the running variable (days relative to birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Significance: * 10%; ** 5%; *** 1%.

4.3 Gender Heterogeneity

There is extensive evidence that women respond to competition differently than men (see Niederle, 2016, for a survey). For example, Gneezy et al. (2003) find that women perform worse on average in competitive environments, but the effect is significantly larger when competing against men. Our setting lets us examine the extent to which there are differential responses by gender to changes in the quality of competition. To study heterogeneity by gender we add an interaction of the treatment indicator with an indicator for male. Thus, the interaction coefficient gives us the difference in treatment effects for males.

Table 5 Panel A shows the effects on quality of assessment groups at age 11 in columns (1) and (2), and at age 13 in columns (3) and (4). On average, the change in the quality of assessment group at age 11 is statistically significantly larger for girls than for boys, although the magnitude of the difference is small (about 4% smaller for boys). The opposite holds at age 13: boys experience a substantially larger change in the quality of competition, as on average, assessment groups are about 10% for boys and 6% faster for girls. These differences between boys and girls reflect the fact that girls tend to mature before boys.

In Panel B of Table 5 we examine effects on own performance. Consistent with the economically small differences in changes in assessment group at age 11, we find no statistically significant differences in the effect on own performance between boys and girls. The point estimate for the interaction term is close to zero and statistically insignificant. At age 13, we find that girls slow down by 0.288% and boys further slow down an additional 0.111%. However, the interaction term is imprecisely estimated when using a cubic function of time in column (4) although similar in magnitude. These small gender differences in performance likely reflect different magnitudes of changes in the assessment groups, rather than reflecting different gender responses to competition.

Panel C presents estimates for the 2SLS regression like in Table 4 but including the interactions with the male indicator. These results suggest that there are no significant differences in gender responses to better assessment groups. Indeed, the point estimates for the interaction terms are close to zero and statistically insignificant. This reinforces that boys and girls respond similar to changes in assessment groups. This is consistent with laboratory evidence showing that there are no gender differences when subjects compete against participants of the same gender (Niederle and Vesterlund, 2007). It is also

Table 5. Effects of Aging-up by Gender

	(1)	(2)	(3)	(4)
	Age 11		Age 13	
*Panel A, Dep. Var.: Log(Average Competitor's time)*100*				
Treatment	-15.44***	-14.81***	-6.183***	-5.725***
	(0.128)	(0.285)	(0.0966)	(0.212)
Treatment*Male	0.689***	0.860*	-3.880***	-3.666***
	(0.212)	(0.479)	(0.170)	(0.369)
*Panel B, Dep. Var.: Log(Swim Time)*100*				
Treatment	0.332***	0.476***	0.288***	0.273***
	(0.0337)	(0.0733)	(0.0284)	(0.0587)
Treatment*Male	0.0650	-0.0314	0.111**	0.145
	(0.0565)	(0.127)	(0.0476)	(0.0987)
*Panel C: 2SLS, Dep. Var.: Log(Swim Time)*100*				
log(Avg. Competition)*100	-0.0216***	-0.0323***	-0.0468***	-0.0478***
	(0.0022)	(0.0050)	(0.0047)	(0.0105)
log(Avg. Competition)*100*Male	-0.0049	0.0005	0.0075	0.0037
	(0.0038)	(0.0089)	(0.0058)	(0.0133)
Obs	441,920	441,920	450,294	450,294
Nr Swimmers	39,750	39,750	45,439	45,439
Controls	Yes	Yes	Yes	Yes
f()	Linear	Cubic	Linear	Cubic

Notes: Sample of swimmers competing within 31 days (both before and after) of their 11^{th} and 13^{th} birthdays. Each coefficient in Panels A and B corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. Coefficients in each column of Panel C correspond to separate 2SLS regression where the instrumented endogenous variable and its interaction with male (the logarithm of the average competitor's race time times 100), are instrumented by the treatment indicator for change in age-group and its interaction with male. Dependent variable in Panel A is the logarithm of the average competitor's times (in seconds) times 100, in Panel B and C it is the logarithm of own swim time (in seconds) times 100. Controls (including their interactions with an indicator for male) include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* in Panels A and B are the treatment indicator, measuring the local average treatment effect of changing age-group. *Male* indicates whether swimmer is male. The functions of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Significance: * 10%; ** 5%; *** 1%.

important to highlight that more girls than boys choose to compete in club swimming competitions. Also, our sample presumably has a stronger preference for competition than the general population. The fact that girls are affected similarly to boys suggests that the findings in the literature could be driven by heterogeneous preferences for competition rather than a gender-specific trait.

4.4 Ability Heterogeneity

A notable advantage of our setting is that we observe each swimmer's prior performance. This lets us construct measures of ex-ante ability to study whether swimmers respond differently across the ability distribution to changes in the quality of the assessment group. This helps us distinguish whether discouragement or rational strategic decisions drive our findings because the likelihood of placing prior to aging up varies significantly across the ability distribution. For example, a median-ability swimmer finished in the top 3 in only 3.5% (2.9%) of races before turning 11 (13) years old. In contrast, the top ability decile finished top 3 in 67% (53.8%) of races before aging up to 11 (or 13). Because median ability swimmers had minimal chances of placing prior to aging up, and high ability swimmers had far larger chances, canonical tournament theory models (e.g., Connelly et al., 2014) would predict that the largest slowdowns would occur at the top of the ability distribution.

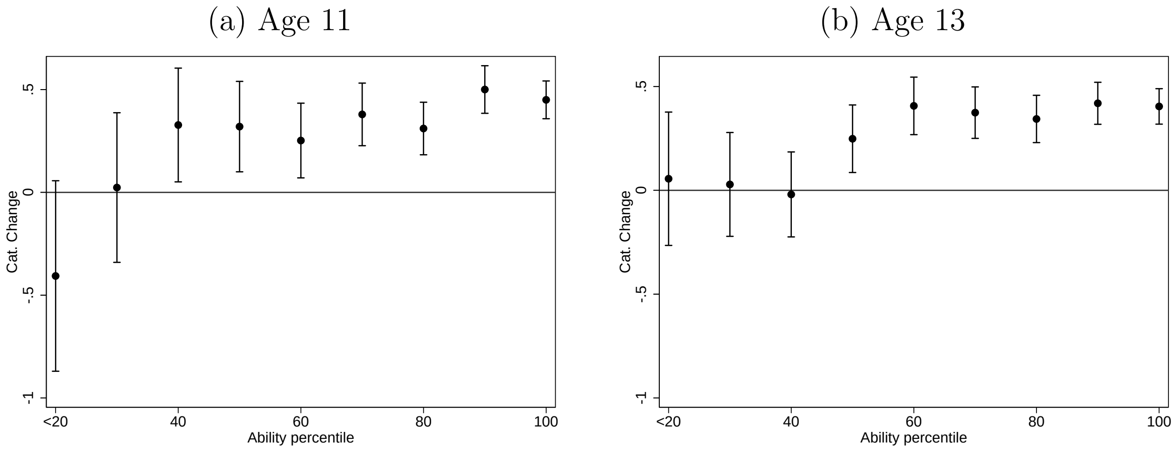
We construct a measure of swimmer's ex-ante ability by using their prior swim history. We identify a swimmer's ability decile by averaging all of the swimmer's Hy-Tek scores (a measure designed to standardize comparisons of times between events and ages) in a six month period before their corresponding birthday. We then estimate equation (2) separately for each decile of ability. The lowest two deciles are aggregated for power.[14]

Figure 3 presents the effects on own performance by ability decile. The figure plots treatment effects coefficients along with the corresponding 95% confidence interval. At the bottom of the distribution (slowest swimmers) we do not find statistically significant effects of changing assessment group. However, after the 40th percentile of the ability distribution we observe positive and statistically significant effects that remain similar in magnitude as we move up the ability distribution, up to the very top. The patterns are similar at the age 11 and 13 cutoffs.

That the magnitude of the slowdown in times upon aging up is unaffected by the large changes in the probability of winning (and hence the marginal impact of effort on winning probabilities) indicates that the slowdown is not driven by the standard predictions of tournament theory that effort should be increasing in the probability of winning.

[14]Each decile is similar in terms of number of swimmers but not in terms of number of races swum. Swimmers in the lower tail of the distribution swim significantly fewer races, thus the smaller sample.

Figure 3. Treatment Effect on Own Time by Swimmer Ability Decile



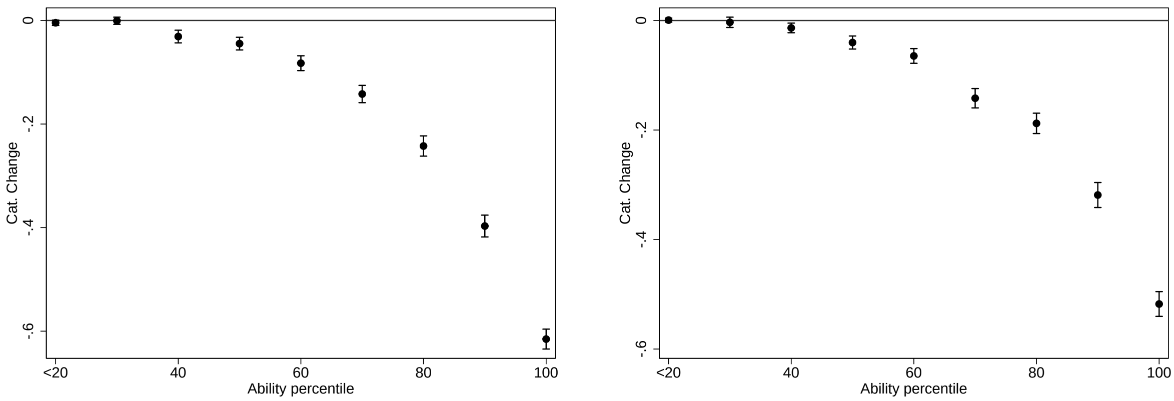
Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each dot corresponds to the treatment indicator of separate OLS regression of equation 2 by ability decile, that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. To calculate a swimmer's ability decile, we first average their Hy-Tek scores across all events and for all times recorded in a window of 2 to 6 months before their Nth birthday. Then we calculate their corresponding decile within the sample of swimmers used for analysis. The lowest two deciles (slowest swimmers) were aggregated. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

To further investigate, we can proxy for the probability of winning by the probability of being *seeded* in the top-3 for that race. We redo the analysis in Figure 4 instead using the probability of being seeded in the top-3 as the main outcome. As we move up the ability distribution, there is a dramatic drop in the probability of still being seeded in the top-3 after aging up. The probability a top swimmer is seeded in the top-3 drops by 60 percentage points at age 11. Therefore, if the slowdown were driven by the decreased (marginal) probability of placement, the effects of changing age-group should be increasing in the ability distribution—mirroring the distribution of effects in seeding. However, the pattern of treatment effects on own performance by ability in Figure 3 is clearly inconsistent with the stark changes in the probability of placing. This indicates that the slowdown in performance is not driven by swimmers being rational agents who reduce effort in response to a reduced probability of winning.

While overall ability can be informative, we can shed further light by exploiting significant heterogeneity *within* a swimmer. In practice, swimmers compete in different strokes and can be strong in one and weak in another. Overall swimmer ability averages across

Figure 4. Treatment effect on *Seeding* in the Top-3 by Swimmer Ability Decile

(a) Age 11 (b) Age 13

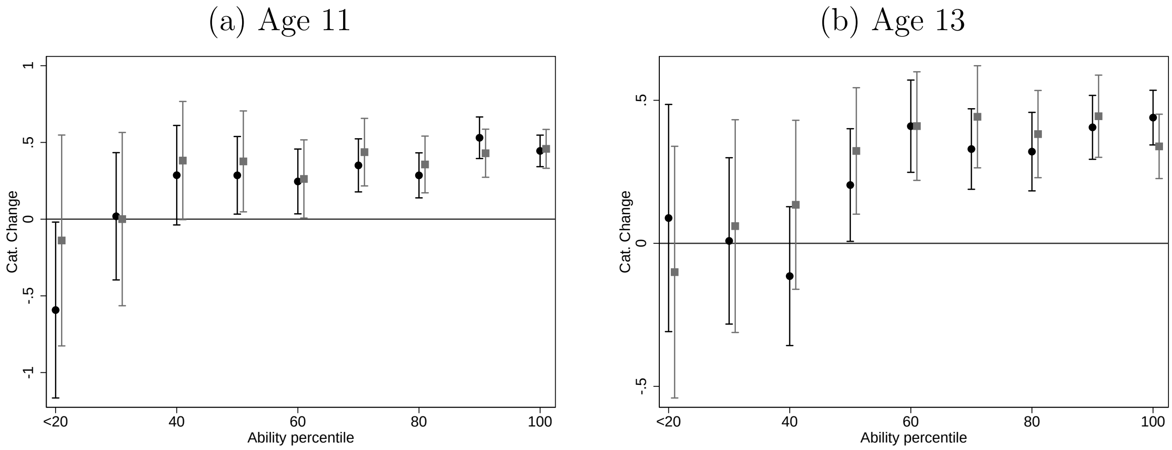


Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is an indicator that equals one if swimmer was *seeded* in the top-3 for that event within corresponding age group. Each dot corresponds to the treatment indicator of separate OLS regression of equation 2 by ability decile, that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. To calculate a swimmer's ability decile, we first average their Hy-Tek scores across all events and for all times recorded in a window of 2 to 6 months before their Nth birthday. Then we calculate their corresponding decile within the sample of swimmers used for analysis. The lowest two deciles (slowest swimmers) were aggregated. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

strong and weak strokes, but the largest changes to the probability of placing at the top matters more for a swimmer's best stroke, so we might be averaging between zero effects and large slow down effects for events in their best stroke. We rule this out by further exploiting differential effects by ability and best stroke.

Figure 5 presents these results, where the grey squares and lines indicate coefficients and 95% confidence intervals for a swimmer's best stroke, and the black coefficients correspond to all other strokes. We do not find robust evidence of differences in responses by best stroke compared to other strokes. Indeed, we fail to reject the equality of effects between best stroke and others. These results further reinforce that the slowdown is not driven by swimmers optimizing their effort as a function of the probability of winning since we would expect to see the strongest effects for their best strokes.

Figure 5. Treatment Effect on Own Time by Swimmer Ability Decile and Best Stroke



Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Grey squares and lines correspond to races in a swimmer's best stroke, the black dot and line correspond to races in all other strokes. Dependent variable is the logarithm of swim time (in seconds) times 100. Each dot corresponds to the treatment indicator of separate OLS regression of equation 2 by ability decile, that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. To calculate a swimmer's ability decile, we first average their Hy-Tek scores across all events and for all times recorded in a window of 2 to 6 months before their Nth birthday. Then we calculate their corresponding decile within the sample of swimmers used for analysis. The lowest two deciles (slowest swimmers) were aggregated. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

4.5 Falsification test: Non age-group related factors

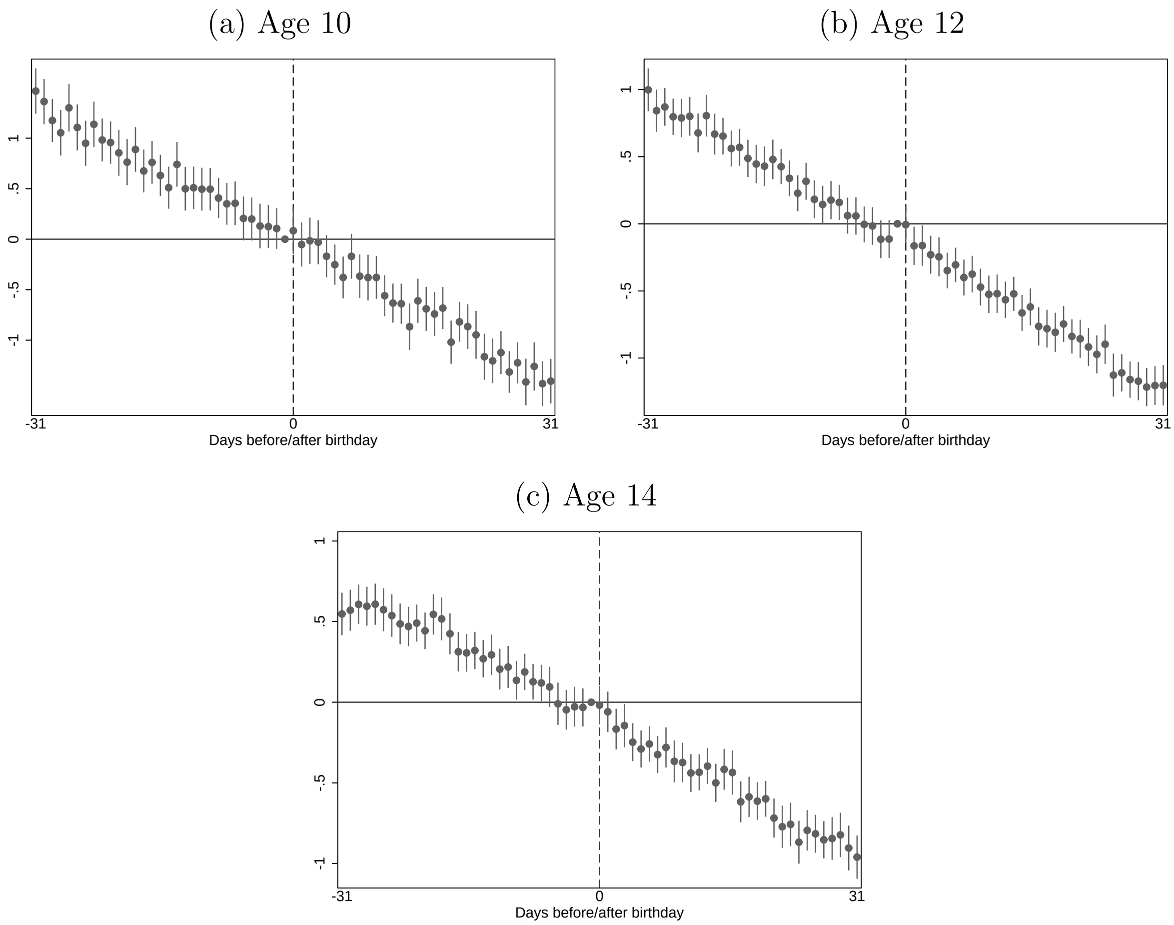
We rely on changing age groups as our source of identification. One concern is that our estimates could be spurious and the product of other events related to a birthday (e.g., celebrations, changes in diet, excitement, distractions, etc.). If these factors are causing swimmers to swim slower then we would expect to find similar effects when examining birthdays where swimmers do not age up.

We replicate our main analysis for placebo birthdays at ages 10, 12 and 14. Figure 6 presents reduced form effects of birthdays on own performance. In stark contrast to Figure 1, there are no visible changes or jumps in own performance upon a birthday with no age group change. Table 6 summarizes the main findings for both competitor's average times and own performance. Panel A shows very small, statistically significant, improvements in the quality of competition (measured as the log of their average time) after a placebo birthday.[15] This largely reflects that as swimmers improve with age they sometimes gain access

[15]See Appendix Figures A.4 for the graphical representation of these estimates.

to competitions with entry time restrictions that, by construction, feature faster cohorts.

Figure 6. Adjusted log(Swim time)*100 relative to placebo birthday



Notes: Sample of swimmers competing in meets within 31 days (before and after) of their 10^{th}, 12^{th} and 14^{th} birthdays. First day of meet used as reference time. Dots are point estimates of running variable dummies (-1 is the omitted category), from an OLS regression that include swimmer-event fixed effects and controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

Panel B of Table 6 presents the effects of non-age-group change birthdays on the logarithm of a swimmer's time (times 100). As Figure 6 suggests, the estimates are close to zero and not statistically significant in most cases. The coefficient for the age 12 placebo is significant at the 5% level, but its magnitude is economically very small—less than one-fifth of the effects we found. Furthermore, these results remain robustly small and statistically insignificant when we use different data windows and sample restrictions (see Appendix Table A4). These results suggest that other events related to a swimmer's birthday do not play a material role in explaining our results for swimmers who age up.

Table 6. Placebo effects on non-age-group change birthdays

	(1)	(2)	(3)	(4)	(5)	(6)
	Age 10		Age 12		Age 14	
Panel A: log(Average Competitor's time) $*$ 100						
Treatment	-0.671***	-0.963***	-0.581***	-0.625***	-0.244***	0.0261
	(0.101)	(0.225)	(0.0725)	(0.160)	(0.0626)	(0.136)
Panel B: log(Swim Time) $*$ 100						
Treatment	0.0401*	0.111	0.0565**	0.0510	-0.0285	-0.0206
	(0.0343)	(0.0758)	(0.0227)	(0.0499)	(0.0205)	(0.0440)
Obs	449,749	449,749	615,613	615,613	584,547	584,547
Nr Swimmers	40,534	40,534	54,223	54,223	51,036	51,036
Controls	Yes	Yes	Yes	Yes	Yes	Yes
f()	Linear	Cubic	Linear	Cubic	Linear	Cubic

Notes: Sample of swimmers competing within 62 days (both before and after) of their 10^{th}, 12^{th} and 14^{th} birthdays. In Panel A, the dependent variable is the logarithm of the average time for all other swimmers in age-group (i.e. 10 and under, 11-12, 13-14) excluding their own time for that event-meet times 100. In Panel B, the dependent variable is the logarithm of own swim time times 100. Each coefficient corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of a birthday (i.e., not changing age group). The functions of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.

5 Conclusion

In this paper we provide evidence that individuals performance worsens on average when assessed relative to better peers. We exploit a natural experiment: sharp age group changes when a swimmer turns 11 or 13 in swimming competitions that strongly improve the set of swimmers against whom the swimmer is being assessed. This setting allows us to circumvent several identification challenges. The effects we find manifests itself in swimmers slowing down significantly upon aging up. Comparable slow-downs are found for both middle- and high-ability swimmers, and for the same swimmer in both their best strokes and others, indicating that swimmers slow down due to psychological discouragement rather than a rational cost-benefit calculation.

There are several psychological theories of motivation and self-assessment that could underlie the discouragement effects we find. For instance, the large changes in the quality of assessment groups upon aging up reduces a swimmer's relative standing—regardless of where they stand in the ability distribution. Lowering self-esteem and self-efficacy has been predicted and shown to lead to discouragement (Bandura, 1977, Vancouver et al., 2002). Feeling competent is also crucial for intrinsic motivation (Deci and Ryan, 1985). When swimmers age up and their relative performance drops, their perceived competence may decrease, leading to reduced intrinsic motivation and increased discouragement.

Our results suggest a discouragement mechanism that may underlie several findings in the education literature where relative assessments are widespread. For instance, students admitted to both more and less selective universities earning similar amounts on average, and the under-performance of both marginally-admitted students attending elite schools and children who start kindergarten younger than their cohort. Our findings may also have implications for the design of incentive schemes. For example, firms using tournaments in contexts with large productivity differences among employees could lead to reduced output relative to less-competitive rewards schemes.

References

Abdulkadiroğlu, A., Angrist, J., and Pathak, P. (2014). The elite illusion: Achievement effects at boston and new york exam schools. *Econometrica*, 82(1):137–196.

Ager, P., Bursztyn, L., Leucht, L., and Voth, H.-J. (2022). Killer incentives: Rivalry, performance and risk-taking among german fighter pilots, 1939–45. *The Review of economic studies*, 89(5):2257–2292.

Allen, E. J., Dechow, P. M., Pope, D. G., and Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63(6):1657–1672.

Arcidiacono, P., Foster, G., Goodpaster, N., and Kinsler, J. (2012). Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics*, 3(3):421–470.

Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics*, 120(3):917–962.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191.

Barankay, I. et al. (2012). Rank incentives: Evidence from a randomized workplace experiment.

Barrow, L., Sartain, L., and De La Torre, M. (2020). Increasing access to selective high schools through place-based affirmative action: Unintended consequences. *American Economic Journal: Applied Economics*, 12(4):135–163.

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.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The quarterly journal of economics*, 117(3):871–915.

Berger, J. and Pope, D. (2011). Can losing lead to winning? *Management Science*, 57(5):817–827.

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

Bottan, N. L. and Perez-Truglia, R. (2022). Choosing your pond: location choices and relative income. *Review of Economics and Statistics*, 104(5):1010–1027.

Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy*, 119(5):982–1013.

Bucciol, A. and Castagnetti, A. (2020). Choking under pressure in archery. *Journal of behavioral and experimental economics*, 89:101581.

Burgess, S., Rawal, S., and Taylor, E. S. (2023). Teachersâ use of class time and student achievement. *Economics of Education Review*, 94:102405.

Burke, M. A. and Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1):51–82.

Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., and Rao, G. (2018). Status goods: experimental evidence from platinum credit cards. *The Quarterly Journal of Economics*, 133(3):1561–1595.

Connelly, B. L., Tihanyi, L., Crook, T. R., and Gangloff, K. A. (2014). Tournament theory: Thirty years of contests and competitions. *Journal of management*, 40(1):16–47.

Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, 107(2):425–56.

Dale, S. B. and Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, 117(4):1491–1527.

Dale, S. B. and Krueger, A. B. (2014). Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of human resources*, 49(2):323–358.

Dasgupta, U., Mani, S., Sharma, S., and Singhal, S. (2022). Effects of peers and rank on cognition, preferences, and personality. *Review of Economics and Statistics*, 104(3):587–601.

de Roux, N. and Riehl, E. (2022). Do college students benefit from placement into higher-achieving classes? *Journal of Public Economics*, 210:104669.

Deci, E. L. and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US, Boston, MA.

Elder, T. E. and Lubotsky, D. H. (2009). Kindergarten entrance age and childrens achievement impacts of state policies, family background, and peers. *Journal of human Resources*, 44(3):641–683.

Fabregas, R. (2023). Trade-offs of attending better schools: Achievement, self-perceptions and educational trajectories. *The Economic Journal*, 133(655):2709–2737.

Figlio, D. and Ozek, U. (2024). The unintended consequences of test-based remediation. *American Economic Journal: Applied Economics*, 16(1):60–89.

Genakos, C. and Pagliero, M. (2012). Interim rank, risk taking, and performance in dynamic tournaments. *Journal of Political Economy*, 120(4):782–813.

Genicot, G. and Ray, D. (2017). Aspirations and inequality. *Econometrica*, 85(2):489–519.

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

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707.

Kuhnen, C. M. and Tymula, A. (2012). Feedback, self-esteem, and performance in organizations. *Management Science*, 58(1):94–113.

Kuziemko, I., Buell, R. W., Reich, T., and Norton, M. I. (2014). Last-place aversion: Evidence and redistributive implications. *The Quarterly Journal of Economics*, 129(1):105–149.

Leuven, E., Oosterbeek, H., Sonnemans, J., and Van Der Klaauw, B. (2011). Incentives versus sorting in tournaments: Evidence from a field experiment. *Journal of Labor Economics*, 29(3):637–658.

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

Matta, R., Ribas, R. P., Sampaio, B., and Sampaio, G. R. (2016). The effect of age at school entry on college admission and earnings: a regression-discontinuity approach. *IZA Journal of Labor Economics*, 5(1):9.

Matthewes, S. H. (2021). Better together? heterogeneous effects of tracking on student achievement. *The Economic Journal*, 131(635):1269–1307.

Mountjoy, J. and Hickman, B. R. (2021). The returns to college(s): Relative value-added and match effects in higher education. Working Paper 29276, National Bureau of Economic Research.

Niederle, M. (2016). Gender. *Handbook of Experimental Economics*, second edition(1):481–553.

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.

Paserman, M. D. (2023). Gender differences in performance in competitive environments? evidence from professional tennis players. *Journal of Economic Behavior Organization*, 212:590–609.

Pope, D. and Simonsohn, U. (2011). Round numbers as goals: Evidence from baseball, sat takers, and the lab. *Psychological science*, 22(1):71–79.

Pope, D. G. and Schweitzer, M. E. (2011). Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1):129–157.

Ribas, R. P., Sampaio, B., and Trevisan, G. (2020). Short-and long-term effects of class assignment: Evidence from a flagship university in brazil. *Labour Economics*, 64:101835.

Rosaz, J., Slonim, R., and Villeval, M. C. (2016). Quitting and peer effects at work. *Labour Economics*, 39:55–67.

Sacerdote, B. et al. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, 3(3):249–277.

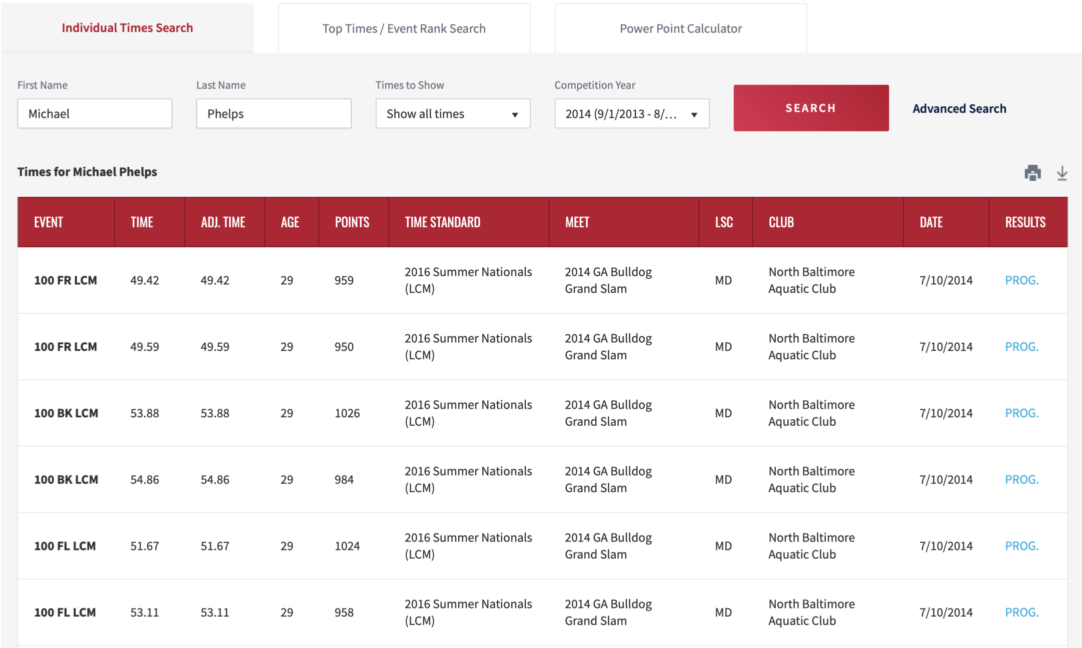
Toma, M. (2017). Missed shots at the free-throw line: Analyzing the determinants of choking under pressure. *Journal of Sports Economics*, 18(6):539–559.

Vancouver, J. B., Thompson, C. M., Tischner, E. C., and Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of applied psychology*, 87(3):506.

Zárate, R. A. (2023). Uncovering peer effects in social and academic skills. *American Economic Journal: Applied Economics*, 15(3):35–79.

A Appendix Figures and Tables

Figure A.1. Screen-shot of USA Swimming Time Search Website

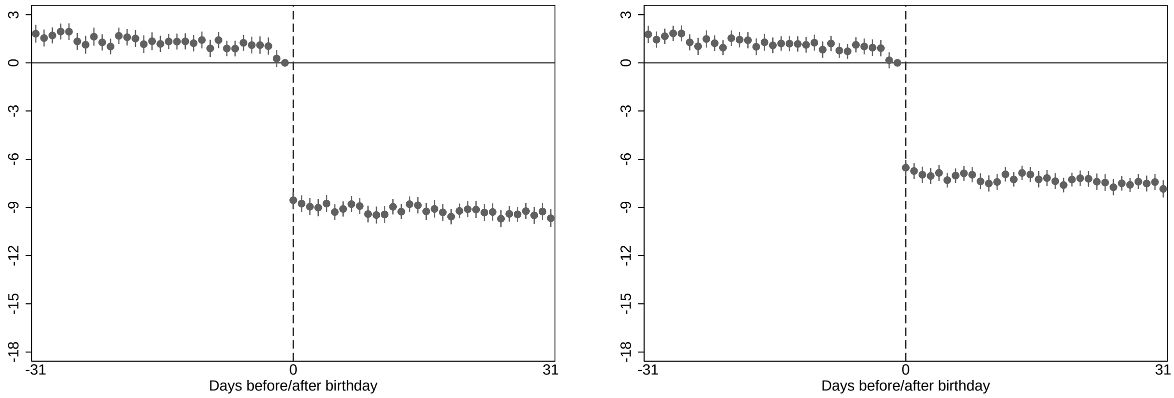


EVENT	TIME	ADJ. TIME	AGE	POINTS	TIME STANDARD	MEET	LSC	CLUB	DATE	RESULTS
100 FR LCM	49.42	49.42	29	959	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.
100 FR LCM	49.59	49.59	29	950	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.
100 BK LCM	53.88	53.88	29	1026	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.
100 BK LCM	54.86	54.86	29	984	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.
100 FL LCM	51.67	51.67	29	1024	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.
100 FL LCM	53.11	53.11	29	958	2016 Summer Nationals (LCM)	2014 GA Bulldog Grand Slam	MD	North Baltimore Aquatic Club	7/10/2014	PROG.

Figure A.2. Adjusted log(Average Competitor time)*100 relative to birthday at age 13 - alternative definition

(a) Alternate (13-16)

(b) Original (13-14)

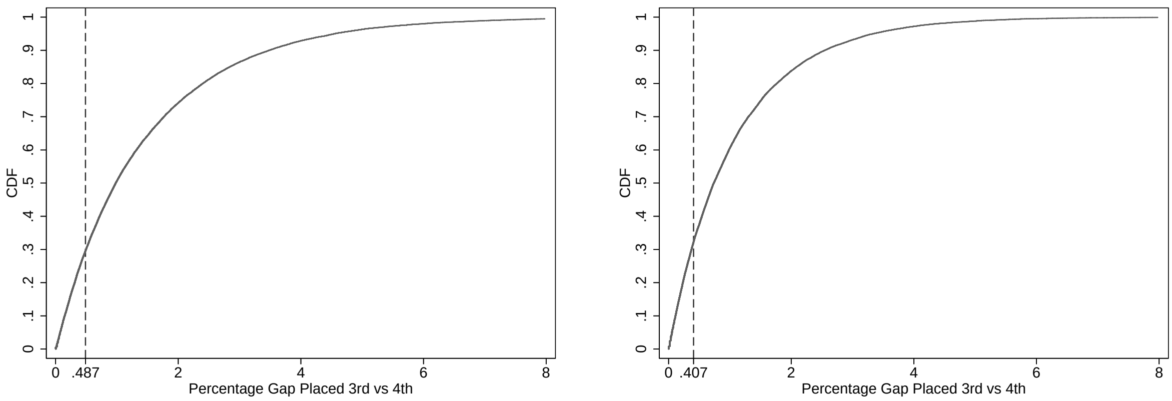


Notes: Sample of swimmers competing within 31 days (before and after) of their 13^{th} birthday, using 13-16 definition for competition after aging up. Dots are point estimates of running variable dummies (-1 is the omitted category) from an OLS regression that include swimmer-event fixed effects and controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

Figure A.3. CDF Percentage Swim Time Gap Between 3rd and 4th Place

(a) 11-12 Age Group

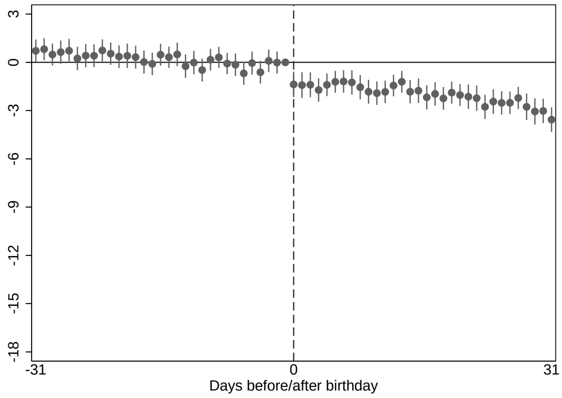
(b) 13-14 Age Group



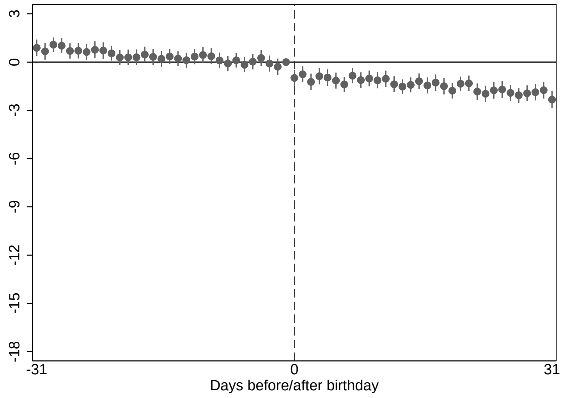
Notes: CDF of percentage time gaps between 3rd and 4th positions in sample of 100 Freestyle Short Course Yard meets with more than 20 participants. Dashed line corresponds to local average treatment effect for aging up to 11 and 13 in panels (a) and (b), respectively, estimated in columns (3) of Appendix Table A5.

Figure A.4. Adjusted log(Average Competitor time)*100 relative to placebo birthday

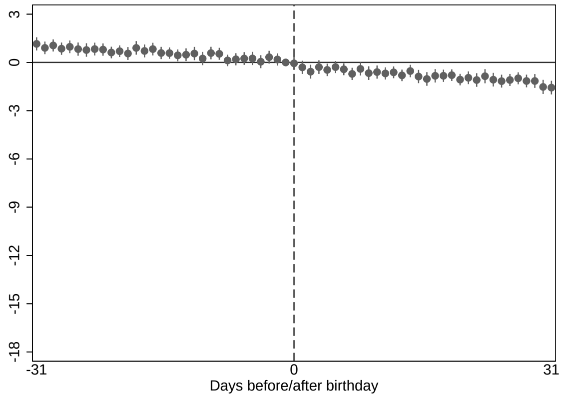
(a) Age 10



(b) Age 12



(c) Age 14



Notes: Sample of swimmers competing within 31 days (before and after) of their 10^{th}, 12^{th} and 14^{th} birthdays. Dots are point estimates of running variable dummies (-1 is the omitted category) from an OLS regression that include swimmer-event fixed effects and controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. Ticks represent 95% confidence intervals, standard errors are clustered at swimmer level.

Table A2. Age grouping of heats for random sample of meets

Percentage of meets grouping:	
Ages 10 and under	41.96%
Ages 9-10	26.79%
Other/Mixed	25.89%
Ages 11-12	**70.27%**
Ages 11 and over	1.34%
Other/Mixed	23.03%
Ages 13-14	**36.61%**
Ages 13 and over	**44.64%**
Other/Mixed	13.39%
Nr Meets	224
Nr LSCs	5

Notes: Sample of invitational meets for the 2015-2016 season for a random sample of Local Swimming Committees (LSCs). Using each meet's information sheet, the authors noted the age cutoffs used for seeding heats for common events (e.g. 100 freestyle). The table shows the percentage of meet that seeded the indicated age group.

Table A3. Sample construction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Age 11 threshold				Age 13 threshold			
	Swims year before birthday	Swimmer-events on both sides	Drop swimmers < 3 meets	Sample in 31 day window	Swims year before birthday	Swimmer-events on both sides	Drop swimmers < 3 meets	Sample in 31 day window
Swimmer Characteristics								
Male	0.38	0.378	0.378	0.378	0.388	0.377	0.369	0.369
	(0.485)	(0.485)	(0.485)	(0.485)	(0.487)	(0.485)	(0.482)	(0.482)
Age Start	8.633	8.493	7.895	7.895	9.727	9.525	9.029	9.029
	(1.288)	(1.314)	(1.078)	(1.078)	(1.758)	(1.767)	(1.513)	(1.513)
Days since last meet	40.121	43.633	43.04	41.171	39.179	42.262	41.248	39.893
	(62.365)	(65.538)	(62.544)	(59.834)	(69.062)	(70.818)	(65.998)	(65.144)
Nr meets	40.538	43.684	42.929	41.883	36.398	37.708	37.761	36.843
	(30.236)	(31.816)	(31.546)	(31.142)	(26.726)	(28.071)	(28.309)	(28.126)
Nr heats	5.425	5.79	6.087	6.025	6.025	6.408	6.565	6.472
	(2.791)	(2.651)	(2.668)	(2.692)	(3.181)	(2.929)	(2.957)	(3.021)
Swim time (in seconds)								
All events	84.693	74.686	75.548	78.246	93.291	99.293	100.454	103.854
	(64.332)	(52.988)	(54.952)	(58.205)	(76.967)	(73.041)	(73.986)	(75.559)
100 Freestyle SCY	78.128	76.742	74.235	74.321	65.973	65.442	64.24	64.277
	(12.325)	(11.796)	(9.620)	(9.787)	(8.523)	(8.157)	(6.832)	(7.031)
Nr swimmers	147,795	55,815	39,750	39,750	144,737	54,703	45,439	45,439
Nr observations	4,097,352	945,098	744,119	441,920	4,432,834	845,278	749,595	450,295

Notes: Standard deviation in parenthesis, within-swimmer standard deviation in brackets.

Table A4. Robustness for placebo effects on non-age-group change birthdays

	(1)	(2)	(3)	(4)	(5)	(6)
*Dep. Var.: log(Swim Time)*100*						
	Panel A: Age 10					
Treatment	0.0764***	0.0291	0.0298	-0.000788	0.0194	0.0330
	(0.0171)	(0.0244)	(0.0315)	(0.0469)	(0.0435)	(0.0326)
Obs	1,220,965	672,614	445,803	337,811	198,369	420,046
Nr Swimmers	62,112	50,701	40,284	34,015	20,823	38,572
	Panel B: Age 12					
Treatment	0.0633***	0.0155	0.0452	0.0575	-0.00549	0.0616**
	(0.0117)	(0.0167)	(0.0238)	(0.0320)	(0.0297)	(0.0226)
Obs	1,665,336	925,208	531,246	469,976	269,226	569,643
Nr Swimmers	81,442	67,387	48,441	46,142	28,194	51,297
	Panel C: Age 14					
Cat Change	-0.00574	-0.0221	-0.0305	-0.0962	-0.0506***	-0.0395
	(0.0111)	(0.0154)	(0.0253)	(0.0261)	(0.0296)	(0.0207)
Obs	1,534,441	870,829	359,881	278,720	449,098	532,304
Nr Swimmers	74,752	62,782	34,825	27,924	43,897	47,755
Window	84	42	31	31	31	31
Other	-	-	No long dist.	Donut	Multi. Event	Regular

Notes: Sample of swimmers competing within given days (both before and after) of their 10^{th}, 12^{th} and 14^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate OLS regression of equation 2 that includes swimmer-event fixed effects. All estimates include controls for: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of changing age-group. All estimates use a linear function of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Window Size* indicates the sample restriction for the number of days before and after a birthday swimmers must compete within. *Other* indicates other sample restrictions. In column (3), *No long dist* drops swimmers who ever participate in events longer than 800 yards or meters. Column (4), *Donut hole* drops observations within 4 days before and after cutoff. Column (5) restricts to swimmers participating in more than 2 meet-events before and after aging up. Column (6), *Regular* drops major championship meets. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.

Table A5. Effects of age-category change by event

	(1) 50 FR SCY	(2) 50 FR LCM	(3) 100 FR SCY	(4) 100 BK SCY	(5) 100 BR SCY	(6) 100 FL SCY	(7) 100 FR LCM	(8) 100 BK LCM
	Panel A: Age 11							
Treatment	0.260*	0.999***	0.487***	0.233	0.544***	0.776***	0.212	0.347
	(0.147)	(0.362)	(0.138)	(0.176)	(0.158)	(0.281)	(0.209)	(0.236)
Obs	51,927	15,211	39,424	25,823	23,589	14,593	14,226	10,779
Nr Swimmers	18,651	5,751	15,503	10,245	9,433	5,701	5,470	4,159
	Panel B: Age 13							
Treatment	0.264**	0.697***	0.407***	0.291**	0.261**	0.375**	0.635***	0.576***
	(0.117)	(0.177)	(0.106)	(0.126)	(0.127)	(0.160)	(0.143)	(0.176)
Obs	52,494	18,875	49,339	37,322	32,232	25,897	20,125	15,995
Nr Swimmers	18,661	6,896	18,508	13,805	12,176	9,668	7,362	5,733

Notes: Sample of swimmers competing within 31 days (before and after) of their 11^{th} and 13^{th} birthdays. Dependent variable is the logarithm of swim time (in seconds) times 100. Each coefficient corresponds to a separate OLS regression of equation 2 by event that includes swimmer-event fixed effects. *FR:* Freestyle, *BK:* Backstroke, *BR:* Breaststroke, *FL:* Fly. *SCY:* Short Course Yards, *LCM:* Long Course Meters. Controls include: day of week dummies, number of heats swum that day, total distance swum, number of long distance swims, and number of days since the last meet. *Treatment* is the treatment indicator, measuring the local average treatment effect of changing age-group. The functions of the running variable (days with respect to Nth birthday) are event-specific and interacted with Treatment. *Linear* indicates that the function used is linear, while *Cubic* indicates that a cubic specification is used. Standard errors are clustered at the swimmer level. Bonferroni adjusted significance: * 10%; ** 5%; *** 1%.