Can Academic Redshirting Shrink the Education Gender Gap? Causal Evidence on Student Achievement and Mental Health

Tímea Laura Molnár
Central European University and IZA*

June 30, 2024

Abstract

I study voluntary delays in the school entry of age-eligible children ("academic redshirting"), using Hungarian administrative data on test scores and medical prescriptions, and survey data on mental health, for 2008–2017. I identify a new Local Average Treatment Effect (LATE) of higher school-starting age (SSA), specifically due to redshirting, by exploiting a unique administrative barrier in Hungary: a school-readiness evaluation, compulsory only for potentially redshirted children born before January 1. By instrumenting with births after January 1, my new LATE estimate of redshirting is for *non-school-ready* children: for whom waiting a year could help overcome developmental deficits, but who are deterred from redshirting by the barrier. I find that higher SSA due to redshirting leads to higher test scores and aspirations, and higher propensity to attend the academic track and to obtain a high school degree, but these positive redshirting effects are only present for *boys*. Redshirted boys are also less anxious, more confident, and less prone to being bullied. To buttress (non-)school-readiness, I show new evidence that boys fall behind already at ages 1–4 in their speaking and attention abilities, and complier boys are indeed the least school-ready. In sum, redshirting sets non-school-ready boys—but not girls—up to be college bound in tracks, and it closes the gender gap in aspirations and in obtaining a high school degree, by helping boys overcome non-school-readiness and improve their student achievement and mental health.

*Central European University (Department of Economics and Business); molnartl@ceu.edu.

This paper is a revised version of the second chapter of my PhD Dissertation, defended at the Vancouver School of Economics (VSE UBC) in 2017. I thank my PhD advisors at the VSE UBC—Kevin Milligan, Nicole Fortin, and Joshua Gottlieb—for their continuous guidance and invaluable advice throughout this project. Also, I am indebted to Catherine Haeck, Zsuzsanna Jován, Gábor Kézdi, Vadim Marmer, Adlai Newson, and Andrea Weber, for numerous suggestions to improve this paper, and to David Card, Pedro Carneiro, Marie Connolly, Dániel Horn, Péter Hudomiet, C. Kirabo Jackson, Brian Jacob, Kevin Lang, Thomas Lemieux, Róbert Lieli, Marit Rehavi, Daniel Shack, and conference and seminar participants for helpful comments. I thank Jiaqi Zou for excellent professional language editing, and to members of the DataBank and the Economics of Education and Labor Economics Research Unit at the Hungarian Institute of Economics, Centre for Economic and Regional Studies, especially János Köllő, Gábor Kertesi, Gábor Kézdi, Zoltán Hermann, Eszter Orosz, Dorina Vincze and Melinda Tir, for providing the datasets of the Hungarian National Assessment of Basic Competences, the Hungarian Life Course Survey, KIR-Stat, and T-Star, and for helping with the data. I thank Gábor Kertesi for providing the municipality-level demographic data set by cohorts of the Hungarian Statistical Authority. The paper has been previously circulated under the titles "*Returns to Starting School Later: Academic Redshirting vs. Lucky Date of Birth*", "*The Impact of Academic Redshirting on Student Outcomes and Mental Health: Evidence from Hungary*", and "*On the Role of Academic Redshirting in Shrinking the Education Gender Gap: Causal Evidence on Student Achievement and Mental Health*". Any remaining errors are solely my own.

1 Introduction

A non-trivial and potentially increasing share of parents are postponing the entrance of their age-eligible children into primary education, a practice known as "academic redshirting;"[1] e.g., 96 percent of 6-year-olds were enrolled in kindergarten in the United States in 1968, a figure which fell to only 84 percent by 2005.[2] The share of redshirted children is 5-10 percent among US kindergartners (Aud et al., 2013; Bassok and Reardon, 2013; Stipek, 2002),[3] with boys and high-status children being over-represented among them (Bassok and Reardon, 2013; Dobkin and Ferreira, 2010; Dhuey et al., 2019; Graue and DiPerna, 2000; Schanzenbach and Larson, 2017; Cook and Kang, 2020), as well as non-school-ready children who typically are the lowest achieving ones (Ricks, 2022; Cook and Kang, 2018). The share of redshirted children is 5-10 percent even in many European countries, where the prescribed school-starting age (SSA) is 6 years of age for all.[4,5] In Hungary, where the prescribed SSA is 6 or 7 based on birth month, the share of redshirted children increases from 6 to 68 percent among students born in September to May—thereby rendering Hungary an ideal setting to study redshirting. With free childcare, (non-)school-readiness is the main driver of redshirting here, and conditional on early childhood development markers, low-status children are not less/more likely to be redshirted.

Academic redshirting is parents' voluntary practice of choosing to delay their children's school entry—particularly of children born before the school enrollment cutoff date—and of consciously opting not to comply with the school entry rule. Despite each year millions of families making decisions about redshirting and its importance for policy, no previous research has identified the causal effect of redshirting itself; i.e., none has identified the effect of higher SSA due to the voluntarily delayed school entry of age-eligible children.[6] While previous research on the effects of higher SSA on compliers with school entry rules and enrollment cutoff dates is deep and extensive, redshirting is a fundamentally different decision made under different considerations for potentially different children on the margin.

[1] Redshirting is "*named for the red jersey worn in intra-team scrimmages by college athletes kept out of competition for a year.*" ("Youngest Kid, Smartest Kid?" by Maria Konnikova (The New Yorker, September 19, 2013)). Also *see* "When Should a Kid Start Kindergarten?" by Elizabeth Weil (The New York Times, June 3, 2007).

[2] According to Deming and Dynarski (2008), the school attendance rate of 6-year-olds has not decreased, but they are more likely to be enrolled in kindergarten rather than first grade. However, only 25 percent of the aforementioned change can be explained by changes in school entry laws that increased kindergarten entrance ages. Since 1965, there has been a trend of moving school enrollment cutoff dates from September-February to June-September in the US (Cannon and Lipscomb, 2008). Note that several states do not have a uniform cutoff date for all school districts (but delegate it to the discretion of Local Education Agencies) and some states do not even regulate kindergarten entry age.

[3] In 2010 in the US, 6 percent of kindergartners entered kindergarten before they turned 5, 85 percent started when they were 5–6 years old, and 9 percent after they turned 6—out of which two-thirds were delayed-entry kindergartners based on school system age requirements ("Kindergarten Entry Status: On-Time, Delayed-Entry, and Repeating Kindergartners," U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011), Preliminary Restricted-Use Data File; available here).

[4] The share of 6-year-old girls and boys who had not yet started (ISCED2011 level 1) primary education in 2010–2018 was, respectively, 6 and 12 percentages in Denmark and Belgium, 7 and 9 in Luxembourg, 6 and 10 in Slovenia, 4 and 5 in Greece, 3 and 4 in Spain, 2 and 2 in Italy, 0.5 and 1.5 in France, and zero in Ireland and the Netherlands (OECD Education Statistics, Education at a glance: Enrollment rate by age (2020 Edition); available here. There may be changing trends in redshirting, as well, given that Denmark's 2007 figures were 11 and 21 percent (Gortz et al., 2018).

[5] European countries also differ in parents' decision-making role in the redshirting process. In countries where, in addition to age, school-readiness is an admission criterion for starting school, parents are not the key decision-makers; the main determinant of SSA is the assessment of either the childcare institution (as in Estonia, Hungary, Liechtenstein, and Luxembourg) or the primary school (as in Austria, Bosnia-Herzegovina, Croatia, the Netherlands, and Serbia) (Key Data on Early Childhood Education and Care (ECEC), 2019, Eurydice Report, EC, pp. 105–107; available here).

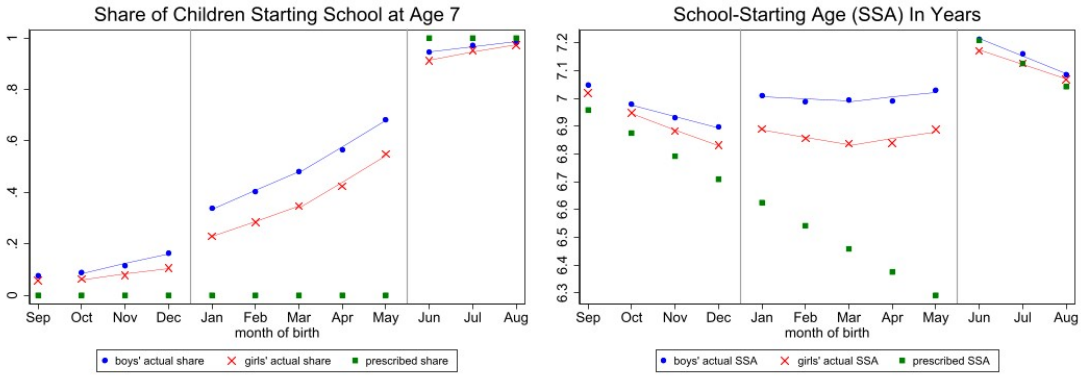
[6] Schanzenbach and Larson (2017) notes that "*[n]o one has conducted a true randomized trial related to redshirting.*"

And, even though Cook and Kang (2018) note that "*[t]here is every reason to believe that redshirting would convey the same "old for grade" benefits,*" it is an empirical question if it indeed does so.

To answer it, I exploit the unique setting in the Hungarian educational system, which has two distinct cutoff dates. First, the primary school enrollment cutoff date is June 1, prescribing children born before June 1 to start school at age 6, and requiring children born after to start school at age 7. Second, the January 1 cutoff affects the practice of redshirting: September-to-May-born children are allowed to be redshirted, but those born before January 1 have to take a school-readiness evaluation, while those born on/after January 1 are exempt from such a test and can be redshirted on the parents' decision alone. I am not aware of such a unique and ideal administrative barrier for identification as the one at the January 1 cutoff in Hungary, elsewhere, which exogenously varies the probability that the parents' intention of redshirting gets realized, independently of their perceptions and preferences.

Figure 1 summarizes how month of birth and actual and prescribed SSA are related in Hungary for 10^{th}-graders in 2008–2017. September-to-May-born children, with prescribed SSA of 6, are allowed to be redshirted, and 6-to-68 percent of them indeed are redshirted and start school at age 7.

Figure 1: **Shares of Children who Start School at Age 7 and School-Starting Age (SSA, Measured in Years of Age), by Month of Birth and Gender, Administrative Data – Grade 10**



My empirical contribution is the novel identification of the Local Average Treatment Effect (LATE) of higher SSA—starting school at 7 *vs.* 6 years of age—that is specifically due to academic redshirting, and providing its first causal estimates. I use an indicator for child birthdays that fall on/after January 1 as an Instrumental Variable (IV), for children born in a 3-months window around that cutoff. I identify the effect of redshirting from the discontinuous jumps at the cutoff, in the share of redshirted children (in the 1^{st}-stage relationship), and in student outcomes (in the reduced-form relationship), for the compliers. While previous research on SSA is deep, direct evidence on redshirting is more relevant for parents (who may consider redshirting, given their children's development before age 6), for policy-makers (who may need to deliberate on regulating parents' opportunities for and institutions' roles in redshirting), and for childcare teachers (who may need to give recommendations on redshirting).

My conceptual contribution is to provide novel insights into how differences in the *complier* population at the January 1 and June 1 cutoffs determine what one can learn from the SSA effect identified from each cutoff. In principle, both higher SSA due to redshirting (the focus of this paper, using the January 1 cutoff), and due to being born after and complying with the enrollment cutoff (the focus of existing literature) are conceptually similar: both lead to a higher SSA and being older in every

grade thereafter. But, the compliers at these cutoffs differ in their underlying school-readiness. Thus, the compliers at the unique cutoff of January 1 in this paper deliver insights about redshirting that is different from what we learn from the typical SSA literature that uses the enrollment cutoff. I show the LATE estimates and the average characteristics of the compliers at both cutoffs.

At the January 1 cutoff, complier children are those who are not redshirted if born before January 1 (in which case the compulsory school-readiness evaluation makes redshirting costly), but who would be redshirted had they been born on or after January 1 (in which case redshirting is relatively costless, as no school-readiness evaluation is required). Many of these complier children are presumably *non-school-ready* (at least perceived by the parents), and for them an extra year before school could help overcome developmental deficits; yet, they are deterred from redshirting by the administrative barrier.

In contrast, in a setting like the Hungarian one, in which parents of children born before the enrollment cutoff have the option to redshirt their younger – and thus less likely to be school-ready – children, compliers at June 1 are *school-ready.*[7] That is, complier children who start school a year older at age 7 do so solely because they were born after the enrollment cutoff of June 1 and, had they been born before, would *not have been redshirted* and started school at a relatively young age of 6.

In sum, compliers at the January 1 cutoff are less likely to be school-ready and are older at a given grade, than the school-ready compliers at June 1 who end up being among the youngest in class.

I show new direct evidence that the compliers are indeed, on average, less likely to be school-ready at the redshirting cutoff of January 1, than at the school enrollment cutoff of June 1. For example, compliers at January 1 are significantly more likely to have been born with low birth weight, been separated from their mother before their 1^{st} birthday, started speaking at an older age, and had problems with speaking, cognition, coordination, and attention at ages 4–5. Within the compliers at January 1, boys started to speak at even older age than girls, and they are more likely to have been diagnosed with ADHD at ages 4–5. Boys are overrepresented among the compliers at the January 1 cutoff, and girls are overrepresented among the compliers at the June 1 cutoff. In Grade 8, compliers at the January 1 cutoff are, on average, more than 1 year older than compliers at the June 1 cutoff. Compliers at the January 1 cutoff are not different from the sample in parental education or income.

I contribute with new descriptive evidence on how the propensity of being redshirted is related to adverse events in early childhood—shock at birth, chronic and developmental difficulties, family tragedies—using exceptionally rich early childhood data. In my setting, while low-status children are more likely to be redshirted, conditional on early childhood developmental markers and shocks low-status children are not less/more likely to be redshirted. I also provide new evidence on developmental gender differences in early childhood, earlier evidence than existing ones from around SSA.[8]

I show that the effects of higher SSA due to academic redshirting (using the January 1 cutoff), *are*, in fact, different from those shown in previous research on compliers with school entry policies and SSA. I show new evidence that the positive redshirting effects among the typically non-school-ready compliers are significant for *boys* only: redshirted boys have better student achievement, and are less anxious, more confident and less prone to being bullied in class. Thus, even though I find boys to fall

[7] In principle, there might also be parents who never want to redshirt, even if their child was non-school-ready and born before the school enrollment cutoff date (say, in May). However, given in Hungary an overwhelming fraction (74 and 60 percent) of May-born boys and girls are redshirted, I assert that the share of such parents is very low.

[8] For instance, Magnuson and Duncan (2016); DiPrete and Jennings (2012); Fidjeland et al. (2023); exception for pre-school ages are, e.g., Brandlistuen et al. (2021) and Stangeland et al. (2018).

behind at ages 1–4 in their speaking, hyperactivity and attention abilities, my results suggest that redshirting helps boys overcome non-school-readiness by improving their student achievement and mental health. For typically young school-ready boys and girls I confirm existing results on positive effects of higher SSA due to complying with the school enrollment cutoff date of June 1.

Given the "boys' crisis" in education (e.g., Fortin et al., 2015; Bertrand and Pan, 2013; Cornwell et al., 2013; Golsteyn and Schils, 2014; Autor et al., 2019), my results that redshirting sets non-school-ready boys—but not girls—up to be college bound in tracks, and that it closes the gender gap in aspirations and in obtaining a high school degree, are striking.[9] Redshirting turns out to be among the very few educational policies that has a bigger effect for boys, and thus shrinks gender gaps.[10]

In what follows, Section 2 reviews the previous literature on SSA, and Section 3 describes the Hungarian ECEC institutional setting, with a focus on the school-readiness evaluation, that is compulsory only for potentially redshirted children born before January 1. Then, Section 4 describes the data sources and outcomes: using Hungarian administrative test score data for 2008–2017, I look at mathematics and reading test scores for Grades 6, 8, and 10, grade repetition, secondary school track choice, and educational aspirations. I link administrative labor records to the test scores, to see actual graduation outcomes from high school, and I link administrative health records to test scores, to look at prescriptions for psycholeptics (a medication for anxiety disorders). Using Hungarian survey data for 2006, I look at self-reported mental health in Grade 8, measured by anxiety and confidence about oneself, and experiences of being bullied in class for one's appearance or student achievement.

Section 5 discusses the identification of the LATE of higher SSA due to academic redshirting: I examine the relationship between SSA and various student achievement and mental health outcomes of children born on either side of the January 1 cutoff, controlling for a linear trend in birth month around the cutoff (with different slopes around), and a rich set of background characteristics. In Section 6, I provide two pieces of reassuring evidence for identification: (i) neither the share of boys nor the share of children of various backgrounds change discontinuously around the January 1 birth cutoff and there is no bunching on either side, and (ii) even though the propensity of being redshirted is related to adverse events in early childhood—such as a shock at birth, developmental difficulties—the fraction of children experiencing such adverse events is not changing discontinuously at the January 1 cutoff.

Section 7 presents the LATE estimates of higher SSA due to academic redshirting: redshirted children generally have better student achievement in secondary school, but these positive redshirting effects are only significant for *boys*: redshirted boys achieve a significantly higher score in mathematics and reading, on average, by 0.2 and 0.25 SDs on Grade 10 testings, respectively, and they are 9.4 percentage points less likely than on-time children to repeat a grade by Grade 10 (which is a large effect size, compared to the baseline share of 11 percent). Moreover, redshirted boys are significantly more likely to (i) attend the academic track (by 11 percentage points), (ii) aspire for a tertiary degree (by 13.4 percentage points, or 26 percent), and (iii) actually obtain a high school degree (by 19 percentage points, or 25 percent), while there is no detectable effect for girls for any of the outcomes. My redshirting estimates are robust to varying windows and functional forms (linear *vs.* quadratic)

[9] Academic redshirting of boys is also the policy recommendation made in the book Reeves (2022).

[10] Haeck et al. (2023) provides an extensive review of the origins of the education gender gap, as well as empirical evidence on early childhood and educational policies, of which neither has a bigger positive impact for boys. Several studies even suggest that girls might benefit more from ECEC enrollment in terms of skill development than boys (e.g., Goodman and Sianesi (2005), Anderson (2008), Cornelissen et al. (2017), Havnes and Mogstad (2015)).

of the trend in month of birth, and to accounting for years spent in childcare and relative age effects.

In comparison, the LATE estimates of higher SSA due to complying with the school enrollment (June 1) cutoff date for mathematics (reading) test scores from Grades 6, 8, and 10 are 0.24, 0.13, and 0.13 (0.28, 0.23, and 0.18) SDs, respectively, which are in line with existing literature (e.g., with Puhani and Weber, 2007; McEwan and Shapiro, 2008; Elder and Lubotsky, 2009); here, boys *and* girls have significantly positive LATE estimates, which is in stark contrast with my LATE estimates of higher SSA due to academic redshirting that are significantly positive for *boys only*.

Finally, Section 8 looks at the mechanisms. I find that academic redshirting helps boys perform better in secondary school because it helps them perform better already in primary school. I also find that redshirted boys are less anxious (and spend less on anxiety medications), are more self-confident and are less prone to being bullied in class. These results suggest that mental health is an important mechanism behind why redshirting improves boys' student achievement: for boys, extra time before school seems to lead to lower anxiety levels, increased self-confidence and better overall well-being.

2 Literature Review

The previous deep research on school-starting age (SSA) overwhelmingly uses school (or kindergarten in case of the US) enrollment cutoff dates or entrance ages for identification, and thus identifies the impact of starting school a year older/later due to being born after and complying with the age cutoff.

In terms of the research design, existing literature typically uses variation in children's SSA that stems from different birthdates and age cutoffs for prescribed school or kindergarten entry, potentially by geography; e.g., Elder and Lubotsky (2009) exploits variation in prescribed kindergarten entry age (KEA) stemming from two sources: the distribution of birthdates among those who comply with their state's enrollment cutoff date, and differences across states in cutoff dates among children born on the same day (but in different states). Similarly, Barua and Lang (2016) instruments actual KEA (6 *vs.* 5) with an indicator of whether the child was required by state law to start school a year older/later (i.e., if the child's birthdate falls after their state's kindergarten entry age cutoff date). For Germany, Puhani and Weber (2007) contrasts children's actual and prescribed SSA, the latter of which depends on birth month and state. Attar and Cohen-Zada (2018) exploits variation in cutoffs that correspond to the same (Jewish lunar calendar) cutoff date in Israel, by comparing the outcomes of children born on different days of the same calendar year over time. Jenkins and Fortner (2019) and Cook and Kang (2020) exploit a policy change in the birthdate enrollment cutoff in North Carolina that forced children born in a six-week window to delay kindergarten entry. These studies—even if using the term 'redshirting'—focus on school entry decisions of those who comply with school enrollment policies.

Existing research focuses primarily on the effect of SSA on student achievement, measured by test scores up to Grade 8, with additional evidence on grade repetition, choice of secondary school track, IQ at age 18, and initial earnings at labor market entry. Non-cognitive and prosocial outcomes are also examined (e.g., persistence, stability, internalizing behavior, teenage pregnancy, and crime).

The literature on test scores is consistent across countries in that children who enter school at an older age—due to the school enrollment cutoff date—perform better on achievement tests; e.g., Altwicker-Hamori and Kollo (2012) provides evidence for Hungary, Attar and Cohen-Zada (2018) for

Israel, Crawford et al. (2010) for England, Dong (2010) and Elder and Lubotsky (2009) for the United States (and Dhuey et al. (2019) for Florida), Fredriksson and Ockert (2014) for Sweden, McEwan and Shapiro (2008) for Chile, Puhani and Weber (2007) for Germany, Ponzo and Scoppa (2014) for Italy, and Smith (2009) for Canada (British Columbia). The range of the impact on test scores is 0.25–0.4 SDs, generally decreasing across Grades 4–8 and slightly larger effects for disadvantaged children; e.g., Datar (2006a) finds that entering kindergarten at an older age is more effective for at-risk children.[11]

Similarly positive effects on grade repetition and secondary school track choice are documented in McEwan and Shapiro (2008), Muhlenweg and Puhani (2010), and Puhani and Weber (2007).[12] Closely related to educational success, Gorlitz et al. (2022) finds that higher SSA increases competencies in receptive vocabulary, and Balestra et al. (2020) finds that children with higher SSA are less likely to develop behavioral and speech problems, but their ADHD and dyslexia/dyscalculia seem unaffected.[13]

Non-cognitive and behavioral measures have also been found to be related to SSA (due to the school enrollment cutoff date); e.g., Cook and Kang (2016) finds that those born just after the kindergarten cutoff date are more likely to drop out of high school and to commit a felony offense by age 19 in the US, while Landerso et al. (2017) finds that a higher SSA lowers crimes committed by boys on both margins in Denmark. Evans et al. (2010) and Elder (2010) show that older children have a significantly lower incidence of ADHD diagnosis and treatment, and Muhlenweg et al. (2012) presents German evidence that a higher SSA has a stable positive effect on persistence, a short-run negative impact on hyperactivity, and a long-run effect on adaptability to change. Fortin et al. (2013) finds that being younger in class aggravates an underlying propensity towards ADHD which is more prevalent for boys in Canada.[14] Dee and Sievertsen (2015) finds that a higher SSA reduces inattention/hyperactivity at ages 7–11 in Denmark, while Datar and Gottfried (2015) finds positive effects on self-control, interpersonal skills and approaches to learning in the US.[15]

Regarding adult outcomes and SSA, the evidence is mixed. Using data on the entire population

[11] Puhani and Weber (2007) presents evidence that entering the German primary school system at age 7 instead of 6 increases Grade 6 reading test scores by 0.4 SDs. Altwicker-Hamori and Kollo (2012) shows that the effects are larger for disadvantaged children. Among children whose mother finished at most primary school, they find very large effects for student achievement in Grade 4 (0.8 SDs) and in Grade 8 (0.25–0.4 SDs). The corresponding effects for children whose mother received a tertiary degree are 0.3 and 0.2 SDs. McEwan and Shapiro (2008) finds that starting school one year later leads to more than a 0.3 SD increase test scores for Grades 4 and 8. They show that entering primary school one year older increases Grade 4 test scores by 0.29 SDs in mathematics and 0.38 SDs in language. Elder and Lubotsky (2009) finds evidence that entering kindergarten one year older leads to a 0.53 SD increase in reading test scores and a 0.83 SD increase in mathematics test scores during the fall semester of kindergarten. However, these effects fade away faster for disadvantaged children. Dhuey et al. (2019) finds that the impact of being one-month older in class (born in September *versus* August) is around 0.2 SDs, irrespective of family background, birth weight, or school quality; moreover, the authors find a positive impact on college attainment and a negative impact on juvenile crime. Attar and Cohen-Zada (2018) finds that entering school a year older boosts Grade 5 test scores in Hebrew by 0.34 SDs and in mathematics by 0.19 SDs; by Grade 8, the effect on mathematics test scores almost doubles.

[12] The authors find that higher SSA decreases (increases) the chance of repeating Grade 1 (in the most advanced school track) by 2 (12) percentage points, while early school entrants are significantly less likely to enter Gymnasium.

[13] At the same time, Barua and Lang (2016) and Fertig and Kluve (2005) do not find any significant effects of higher KEA and SSA on educational attainment, respectively, and Hemelt and Rosen (2016) find mixed evidence that children eligible to start school at a younger age are more likely to complete high school, but underperform while enrolled.

[14] Exploiting US state-level variation in school enrollment cutoff dates, McAdams (2016) uses a Difference-in-Difference identification strategy, and finds that starting school older reduces the incidence of incarceration.

[15] Yet, Lubotsky and Kaestner (2016) does not find that internalizing behavior or self-control were related to KEA, and Pena and Duckworth (2018) finds incosistent effects of SSA on grit.

of Norway, Black et al. (2011) finds a small positive effect of lower SSA on IQ scores at age 18, and a short-run positive effect on earnings that disappears by age 30. Using data from California and Texas, Dobkin and Ferreira (2010) finds no evidence that SSA would affect job market outcomes, such as wages or the probability of employment. For Sweden, Fredriksson and Ockert (2014) finds that SSA affects educational attainment and the timing of labor supply but not prime-age earnings (except for low-status). Du et al. (2012) finds suggestive evidence that relative age has a long-lasting influence on career, given the disproportionately small share of summer-born CEOs. Arnold and Depew (2018) finds that a higher SSA for boys is linked to better health in adulthood, but with no long-term labor impacts. Oosterbeek et al. (2021) finds that children born just before the school enrollment cutoff enter the labor market younger and with more experience, leading to higher earnings until age 40.

Within the literature on SSA, relative age (RA) effects have been found to be positively related to educational success. For example, Dhuey and Lipscomb (2010) shows that an additional month in RA decreases the probability of receiving special education by 2–5 percent, and Bedard and Dhuey (2006), Pena (2017), and Sprietsma (2010) show that it increases test scores. Dhuey and Lipscomb (2008) finds that the oldest students are more likely to be leaders in high school (and enjoy the associated wage premium), and Schneeweis and Zweimuller (2014) shows that they choose more ambitious secondary tracks. Patalay et al. (2015) finds that lower RA is associated with increased internalizing symptoms, poorer peer relationships, and higher impact of mental health difficulties on functioning at school and home, and Aliprantis (2014) finds that returns to RA may be negative for the oldest students.[16]

Existing literature on relative age effects does not definitely answer the question of whether these effects remain important beyond school-starting age, which may be an important consideration for parents (who need to decide whether to redshirt their children or be concerned about whether other parents' redshirting behavior would harm their children) and policy-makers (for how to regulate school-starting age). Cascio and Schanzenbach (2016) supplies evidence from the Project STAR experiment, in which kindergarten-aged children were randomly assigned to classrooms, effectively randomizing their relative age. They find that holding own age constant, younger children performed no worse on achievement tests, and were no more (less) likely to repeat a grade (take college-entry exams).

Evidence on selection into academic redshirting is more scarce, beyond the result that, at least in the US, it is more common among males, whites and high-status non-minority children (e.g., Bassok and Reardon, 2013; Schanzenbach and Larson, 2017; Dobkin and Ferreira, 2010). The results of Cook and Kang (2018) suggest an interaction between abilities and family background: among first-graders in North Carolina public schools, low-status children are more likely to be redshirted without controlling for academic indicators capturing school-readiness, but are less likely to be redshirted once those are taken into account. Jenkins and Fortner (2017) finds that redshirters are more prone to have special educational needs, but also a positive selection into redshirting. On data on kindergartners at Michigan public schools and a Marginal Treatment Effects (MTE) framework, Ricks (2022) finds that redshirters would have been the lowest-achievers otherwise, but they benefit the most from it. Finally, some studies also analyze implications on gaps and equity-efficiency trade-offs (Ricks, 2022; Lenard and Pena, 2018), with some focusing on gender gaps (e.g., Cook and Kang (2018, 2020)).

[16] Johansen (2021) finds that the youngest girls in class are more likely to have abortion, alcohol poisoning, and earlier motherhood (Black et al. (2011) substantiates these for Norway and teenage pregnancy), Ballatore et al. (2020) finds that they are more likely to being victimized (bullied), and Fumarco et al. (2024) finds that RA affects eating behavior.

3 Institutional Background

Hungary has a universal and free childcare system, in which more than 90 percent of children spend at least 3 years in childcare before primary school.[17] The last childcare year is the school-preparation year, roughly equivalent to the North-American kindergarten year. The minimum compulsory childcare is 4 hours per day, starting from September 1 of the calendar year in which the child turns 5 (Public Education Act, 24(3)). Then, academic redshirting is the voluntary delay of school entry of an age-eligible child, which, in the Hungarian context, means postponing the start of primary school to the next academic year and repeating the school-preparation year (which, in the North-American context, would be equivalent to repeating the 'developmental kindergarten' or 'young kindergarten'). Given the free childcare, there is no direct cost of redshirting in Hungary (differently from North America).

In Hungary, the compulsory childcare ends with the start of compulsory schooling. All children who turn 6 before June 1 in a given calendar year are prescribed to start primary school on September 1 of that particular calendar year (Public Education Act 6(1)). Thus, the Hungarian school enrollment cutoff date is June 1. Then, if complying with the prescribed SSA, (i) children born between January and May start primary school at an average age of 6 years 7.5 months to 6 years 3.5 months, respectively; (ii) children born between June and August start primary school at an average age of 7 years 2.5 months to 7 years 0.5 months; and (iii) children born between September and December start primary school at an average age of 6 years 8.5 months to 6 years 11.5 months. This rule generally results in May-born children being the youngest, and June-born children being the oldest in a class.[18]

With some restrictions, redshirting is possible in Hungary, and before 2018 it was decided by the parents together with the childcare institution's board of teachers, and (depending birth month) the local so-called 'Developmental Advisory Board' (DAB). While both the childcare institution and the local DAB are part of the public education system, the latter is an independent pedagogic institution maintained by local governments, which evaluates the school-readiness of potentially redshirted children (Hungarian Public Education Act, 35(4)). The local DAB also evaluates children's learning and socialization skills until adulthood, and proposes skill-enhancing activities involving the parents.

Table 1: **Predicted Path into Primary School, by Month of Birth**

Month of Birth:	Prescribed School-Starting Age (SSA)	Opportunity to Redshirt?	Regime
September–December	6 years, 8.5–11.5 months	yes, with administrative barrier	1
January–May	6 years, 3.5–7.5 months	yes, no administrative barrier	2
June–August	7 years, 0.5–2.5 months	no	3

Table 1 summarizes primary school entry in Hungary, by month of birth, as they pertained to before 2018. *Regime1* pertains to children born in the fall term (September–December): these children are prescribed to enter primary school at age 6, and the administrative barrier to academic redshirting is present for them—they can be redshirted only if the parents request an additional year

[17] For more institutional details, see Appendix A.

[18] While, under exceptional circumstances, it is possible for a child to start school a year earlier than prescribed, this is rare: roughly 0.3 (0.7) percent of children born January–May (September–December) start school at age 5. 94, 97, and 98 percent of those born in June, July, and August, respectively, start school at age 7 (as prescribed).

in childcare and have their children take the compulsory school-readiness evaluation, and both the childcare institution's board of teachers and the local DAB support the parents' request.[19] *Regime2* pertains to children born in the spring term (January–May): these children are prescribed to enter primary school at age 6, and there is no administrative barrier to redshirting for them—they can be redshirted if the parents request an additional year in childcare, and the childcare institution's board of teachers supports the parents' request,[20]; i.e., the local DAB's evaluation or support is not needed in this case. *Regime3* pertains to children born in the summer term (June–August): these children are already required to enter primary school at age 7, and thus redshirting is not allowed for them.

Besides SSA, the January 1 cutoff also affects childcare-starting age (CSA), although to a much smaller extent. Children born in September–to–December are more likely to enter childcare early at age 2, as these children are preferred by the childcare institutions if there is a waitlist;[21] these children will have already spent 4 years in childcare when the possibility of redshirting becomes relevant.

The school-readiness examination conducted by local DABs is standardized and free of charge. Developmental experts first examine how children behave in social settings with other children – typically, during drawing exercises, they investigate to what extent children are able to work while being distracted, able to adapt, and able to cope with new situations, and how much they are influenced by others. Then, in a one-on-one situation, the developmental experts assess the extent of children's awareness about their environment and themselves, as well as their cognitive and recall abilities.

For insight into the costs of a school-readiness examination from the perspective of the parents, I conducted interviews with experts of local DABs in Budapest, Hungary, in the summer of 2013. First, in most cases, the accompanying parent needs to miss work for the duration of the examination. Additionally, travel costs are incurred by families that live in municipalities that do not have a local DAB (for context, there are 4-5 local DABs in each of Hungary's 20 counties). Second, parents might be afraid of their children potentially receiving a poor or problematic evaluation, the documentation of which parents may presume to harm their children's educational career later. Third, parents might be reluctant to go in front of an unknown formal committee, especially if the default option is to discuss their children's development and school entry with the childcare teachers, with whom the parents are already familiar. But, since childcare is free in Hungary, there are no other direct costs of redshirting.

In sum, the main institutional reason for expecting a discontinuous jump in the share of redshirted children at the January 1 cutoff is the higher cost of redshirting a child born before January 1, due to the compulsory school-readiness evaluation. I use this source of exogenous institutional variation at the January 1 cutoff to identify the LATE of higher SSA due to academic redshirting specifically. In Appendix H, I show that my main estimates of the LATE of redshirting are remarkably robust to accounting for the potentially endogenous nature of CSA and years spent in childcare ($yearsCC$).

[19] Hungarian Public Education Act, 24(5).

[20] According to the enactment of the Ministry of Education, 11/1994. (VI. 8.), 15(5)b, the childcare institution's board is obliged to suggest an additional year in childcare for children assigned to start primary school on September 1 in a given AY, but who are not school-ready based on the opinion of the board of childcare teachers.

[21] Hungarian Public Education Act 24(1): childcare institutions can admit 2-year-old children, provided that all the 3-year-old and older applicants living in the municipality are admitted. This being said, children who are turning 3 within 6 months of acceptance are preferred. As acceptances typically take place at the beginning of June and childcare starts only at the beginning of September, children born in January–to–May are most likely to enter childcare at age 3, while those born in July–to–August at age 4.

4 Data, Sample, and Measurement

4.1 Sources: Administrative Data on Test Scores, Prescriptions and High School Graduation, and Survey Data on Mental Health and Early Childhood Events

I use (1) administrative data from the *Hungarian National Assessment of Basic Competences* (HNABC), spanning 2008–2017, for Grades 6, 8, and 10, which includes student test scores in mathematics and reading, grade repetition, secondary school track choice, educational aspirations, year and month of birth, childcare attendance, and background characteristics; (2) survey data from the *Hungarian Life Course Survey* (HLCS) on self-reported mental health indicators, such as anxiety and mental exhaustion, confidence and experiences of being bullied in class, and early childhood events for 8^{th}-graders in 2006; and (3) administrative data (for a 50 percent random sample of (1)), spanning 2009–2017, on medical prescriptions for psycholeptics (treating anxiety), and on high school graduation.[22]

The HNABC does not assess students' factual knowledge of the compulsory curriculum. Instead, the tests seek to measure the extent to which students are able to apply their acquired skills in realistic settings, and the extent to which they possess the necessary competences for further development. The tests are low-stake, the completion of the tests are mandatory for all children, almost all students' tests are centrally processed, and the results count toward each school's average.[23]

I obtain student background variables, year and month of birth, and time spent in childcare from the HNABC student background survey (exact day of birth is not available). This survey provides detailed information about the students' demographic and family background, but its completion was non-compulsory. I restrict the analysis to children for whom information about gender, time spent in childcare, and year and month of birth is available and who have at least one valid test score.[24]

The HLCS is an individual panel survey conducted annually. The original 2006 sample (with 10,022 respondents) is a selected 10 percent random sample from the population of 8^{th}-graders with valid HNABC test scores in that year. Students were followed throughout their school career and were asked questions on their ethnicity, school, and family background (including poverty and home environment). Each wave had a particular block of questions on, for instance, early childhood environment, secondary

[22]The administrative data used in this paper are the property of the *National Health Insurance Fund Administration*, the *Central Administration of National Pension Insurance*, the *National Tax and Customs Administration*, the *National Employment Service*, and the *Educational Authority of Hungary*. The data were processed and provided by the *Databank of the Centre for Economic and Regional Studies in Hungary*.

[23]Exceptions are made for students with special educational needs or autistic students. While students who suffer from behavioral problems or dyslexia/dysgraphia/dyscalculia are required to complete the test sheet, their results are not processed centrally. Children who suffered a temporary injury that rendered them physically unable to complete the test were exempted, as were those who missed class on the day of the test.

[24]The final sample contains 80-83 percent of the original sample (see Table C1). Reassuringly, the excluded observations seem to be random by gender, as there is no systematic relationship found between gender and non-response in the background survey, in the missing response analysis of the 2006–2007 HNABC of the Economics of Education and Labor Economics Research Unit at the *Hungarian Institute of Economics, Centre for Economic and Regional Studies* (available here, in Hungarian). While there is no information available about non-response and birth date, this analysis suggests that students with a privileged family background and better student achievement may be over-represented in my sample; it finds that among students with no valid test scores, those with lower parental education are overrepresented, and that students with valid test scores obtained a systematically higher grade from mathematics the year before. Given that only the effect of higher SSA due to the June 1 enrollment cutoff (but not that of higher SSA due to redshirting) is found to be slightly higher for disadvantaged children (which results are available upon request), the estimated effects shown in this paper are likely a lower bound, and the true impacts are even higher in absolute value.

school application, alcohol and drug usage, social networks, and prejudices.

Administrative data on medication expenditures for a 50 percent random sample of the HNABC, spanning the years 2009–2017, is merged from the *National Health Insurance Fund Administration*. The data contain all social insurance expenditures on medications prescribed either by a public or private provider, by codes of "Anatomical Therapeutic Chemical" (ATC). For the same subsample, I also observe the highest completed secondary-level education (but not yet tertiary-level studies).

4.2 Measurement of Outcomes, Endogenous Variables, and Control Variables

With respect to students' performance in school, I measure:
— *student achievement* by standardized test scores which have mean 0 and standard deviation of 1, by the interactions of grade (6/8/10), year (2008-2017), and subject (mathematics/reading);
— *grade repetition* by a binary variable indicating repeating at least one grade by Grades 6, 8, or 10;
— *secondary school track choice* by binary variables indicating whether the student attends the j-th track in Grade 10, where track #1 is the lowest secondary school track, granting a vocational degree upon completion but no high school diploma, and continuation onto tertiary education is not allowed; track #2 is the middle secondary school track, granting both vocational and high school diplomas upon completion; and track #3 is the highest (academic) secondary school track, which grants a high school diploma upon completion; #2 and #3 allow continuation onto tertiary education;
— *educational aspirations* by indicators of whether the student aspires to earn a high school diploma or more advanced post-secondary degree(s), as their highest level of education;
— *graduation* by indicators of whether the student drops out of secondary school without a degree, or obtains a vocational degree or a secondary school (in other words high school) degree.

With respect to the mental health and student well-being outcomes, I measure:
— *standardized anxiety* by the (standardized) mean score from students' responses to how often they experience the following indicators of Generalized Anxiety Disorder (GAD) (score 5: 'daily'; 4: 'several times per week'; 3: 'weekly'; 2: 'monthly'; 1: 'never or rarely'): (i) having stomachache, (ii) being in bad mood and fatigued, (iii) being irritable, (iv) being fearful, (v) being nervous, (vi) having problems with falling asleep, (vii) waking up during the night, (viii) being dizzy, (ix) being exhausted, (x) feeling nausea and vomiting; then, *often anxious* is a binary variable indicating whether the student experiences any of the above (i)–(x) conditions daily or several times per week;
— *lack of confidence* by a binary variable indicating whether the student agrees with any of the followings: (i) "*I am inclined to consider myself unsuccessful without any talent.*"; (ii) "*I feel I cannot be proud of myself.*"; (iii) "*I wish I had more respect for myself.*"; (iv) "*I feel I am not good at anything.*";
— *bullied* by a binary variable for being bullied in class for appearance or poor student achievement;
— *N05d expenditures* which capture prescriptions for psycholeptics (ATC code N05; antipsychotics and sedatives form the major groups in this category, primarily for treating anxiety).

With respect to background variables, the background survey of the HNABC does not contain information on whether the child was evaluated by the school-readiness committee or whether concerns were expressed about their school-readiness around the age of 6. The exact day of birth is also not available, which would be needed for the implementation of an RDD by day of birth. I construct SSA

from birth date, time spent in childcare, and grade repetition as follows. First, the child's age at the time of the test, in absence of exact birth dates, can be estimated by assuming that all children were born on the 15^{th} of their respective birth month and that the test was taken on May 15^{th}:

$$\text{age-at-the-test (in months)} = (\text{year-of-the-test} - \text{year-of-birth} - 1) \times 12 + (12.5 - \text{month-of-birth}) + 4.5.$$

Second, SSA can be backed out with age-at-the-test and subtracting time elapsed between primary school entry and the Grade 6 test. Assuming that primary school starts on September 1 in all years, students who did not repeat any grade spent 68.5 months between primary school entry and the Grade 6 test (and the same logic follows for Grades 8 and 10). Similarly, with repeating 1, 2 and 3 grades, the time between primary school entry and Grade 6 becomes 80.5, 92.5, and 104.5 months, respectively. I verify that SSA is consistent with information on childcare-starting age and years in childcare.

I measure parental education with four categories: at most primary education, obtained a vocational degree, a secondary school degree, or a tertiary degree. I label children as 'disadvantaged' or 'low-status' if their parents obtained at most a vocational degree, but not a high school degree. The child background control variables are indicators about the family's financial situation (e.g., whether the child is considered to be disadvantaged or excessively disadvantaged by law (meaning being eligible for child protection support, or, having, in addition, parents who completed at most Grade 8 (Act No. LXXIX of 1993 on Public Education (121(1)), whether the child considers their family to be poor, whether the child is eligible for free meals and/or free books at school, whether the child has their own desk at home), the composition of the household (e.g., members of the household, whether the child lives together with a biological father and mother *vs.* a stepfather and stepmother), goods at home (internet, books), and the characteristics of the municipality (region and type of settlement).

The Hungarian Life Course Survey includes very rich data on early childhood events, starting from the birth of the child, including information on family shocks at various ages (such as income shocks and separation from parents), chronic conditions (such as ear infections or asthma) and various developmental markers (such as age of starting to speak or being diagnosed with ADHD at ages 4–5).

5 Identification

In this section, I present the identification of the Local Average Treatment Effect (LATE) of higher school-starting age (SSA) (1) due to academic redshirting, using the January 1 cutoff, and (2) due to complying with the school enrollment cutoff of June 1. I capture SSA with an indicator variable of starting school at age 7 (*vs.* 6). In my main identification strategy (1), I use being born on/after January 1 as an Instrumental Variable (IV) for starting school at age 7, and compare academic achievement and mental health outcomes of children born on either side of the January 1 cutoff in a 3-month window, controlling for a linear trend in birth month around the cutoff and a rich set of background characteristics; I allow for the linear trend to have different slopes on the two sides of the cutoff. After the estimation equations, the IV validity requirements, and the robustness checks, I discuss who the likely compliers are in my research design, and how to estimate their average characteristics. Finally, I discuss the interpretation of my LATE estimates in light of the existing literature on SSA.

5.1 The Identification Challenge and the Intuition Behind Identification

Identification of the effect of starting school at age 7 (instead of age 6) is difficult due to the fact that, in the Hungarian educational system, as in many others, primary SSA is determined not only by an exogenous rule based on date of birth, but also by children's abilities at their entry-relevant age and/or their parents' preferences—both of which are unobserved in my data. As a consequence, Ordinary Least Squares (OLS) estimates of the effect of SSA on student achievement may be biased.

Unobserved abilities bias the OLS estimates downward. This is because unobserved child abilities around age 6 are negatively correlated with their propensity to start school a year later. Furthermore, all else equal, a child with worse abilities has, on average, worse student achievement.

Unobserved parental preferences, on the other hand, lead to a bias in the OLS estimates of ambiguous sign, as it depends on the characteristics of the parents with a higher valuation of childcare services, who are more likely to delay their children's primary school entry. On the one hand, these parents may be those who try to compensate for insufficient skill-enhancing resources at home, and whose children presumably score lower on achievement tests, *ceteris paribus*—leading to a downward bias. On the other hand, these parents may be those who, besides other early childhood resources, view childcare as equally important and want to provide it in the greatest possible extent, and whose children presumably score higher on achievement tests, *ceteris paribus*—leading to an upward bias.

One option to eliminate the bias is to use within-family variation across siblings, as in Dhuey et al. (2019). As family identifiers are unavailable in my data, I use the Wald estimand in an IV design.

The idea behind my identification strategy is shown in Figures 1–2 for 10^{th}-graders in 2008–2017. In Figure 1, the raw first-stage relationship between birth month and the share of children with SSA=7 is shown on the left, and the relationship between birth month and (actual and prescribed) SSA is shown on the right. The cutoffs of January 1 and June 1 are marked with solid lines. Figure 2 shows the raw reduced-form relationship between month of birth and student outcomes in Grade 10.

In Figure 1, the share of Grade 10 children who start school at age 7 steadily increases over birth months September–August, but there are clear jumps at the cutoffs. At the January 1 cutoff, the share of redshirted boys sharply increases by 17.5 percentage points (from 16.5 to 34), and at the June 1 cutoff, the share of boys starting school at age 7 sharply increases by 26 percentage points (from 68 to 94). These jumps at the January 1 and June 1 cutoffs are used as instruments for $SSA = 7$ due to academic redshirting and due to the school enrollment cutoff, respectively. Intuitively, the estimate for the LATE of academic redshirting is the reduced-form estimate rescaled by the first-stage estimate, controlling for a linear trend on the two sides of the January 1 birth date and background variables.[25]

5.2 The IV Design and the First-Stage & Second-Stage Equations

To overcome the bias in the OLS estimates, I use the two discontinuity cutoffs in birth months, (1) January 1 and (2) June 1, to identify the LATE of higher SSA due to (1) academic redshirting, and (2) the school enrollment cutoff date, respectively. I use the Wald estimand for an IV design that, under

[25] The average test scores of boys born in December *vs.* January increases by 1.9 and 1.4 percentage points, respectively, and the share of (i) grade repeaters by Grade 10, (ii) those who aspire for a $HS+$ (tertiary) degree, and (iii) those who end up in the vocational and academic high school track, increase by -0.6, 0.7, -0.9, and 0.6 percentage points.

the conditions discussed below, captures the LATE on *compliers*, who, by definition, are children whose SSA changes, if hypothetically moving from the left of the cutoff to the right.

As day of birth is not available in my data, the bins on either side of the cutoffs contain children of the same birth month. I use a 3-month window, leading to 3-3 bins, by using data on children born in (1) October–to–March, and (2) March–to–August. I control for month of birth linearly, and I allow for this linear trend to have different slopes on the two sides of the cutoffs.

Formally, the Two-stage Least Squares (2SLS) first-stage and second-stage equations are:

$$D_i = \beta_0 + \beta_1 1\{X_i \geq x_d\} + \beta_2 X_i + \beta_3 1\{X_i \geq x_d\} \times X_i + \beta_4 C_i + \sum_{t=1}^{T-1} \mu_{1t} F_{ti} + \eta_i, \tag{1}$$

$$Y_i = \gamma_0 + \gamma_1 D_i + \gamma_2 X_i + \gamma_3 1\{X_i \geq x_d\} \times X_i + \gamma_4 C_i + \sum_{t=1}^{T-1} \omega_{1t} F_{ti} + \upsilon_i, \tag{2}$$

where Y_i is the outcome variable from the set of {test scores, grade repetition, secondary school track choice, aspirations, obtaining a high school degree, mental well-being}; D_i is a binary variable denoting age at primary school start of child i (1: child started primary school at age 7, or $SSA_i = 7$; 0: child started primary school at age 6, or $SSA_i = 6$); X_i is month of birth (linear trend, re-centered at x_d), where the discontinuity cutoff x_d is either January 1 or June 1, and $1\{X_i \geq x_d\}$ is the discontinuity dummy; C_i denotes the vector of control variables, and F_t a full set of birth cohort dummies (defined to start in September and end in August). Only the discontinuity dummy $1\{X_i \geq x_d\}$ is used as an instrument, and is excluded from the second-stage eq. (2). The coefficients of interest are β_1 and γ_1.

5.3 Discussion of IV Validity and Robustness Checks

It is implausible that the potential outcomes, and thus the effect of school-starting age, are homogeneous in the population. Therefore, following Angrist and Pischke (2008), the requirements for the instrument $1\{X \geq x_d\}$ to be valid in a heterogeneous framework are that (i) the instrument is related to the endogenous variable (*relevance, or the existence of a strong first-stage relationship*); (ii) the instrument is as good as randomly assigned: it is independent of the vector of potential outcomes and potential treatment assignments (*independence*); (iii) the instrument has no direct relationship to student achievement, only an indirect relationship through the decision made about primary school-starting age (*exclusion restriction*). In turn, I discuss each of these requirements in detail.

The *first-stage* requires birth month to be systematically and strongly related to school-starting age: children who were born on or just after January 1 and June 1 should have significantly higher probability of starting primary school at age 7 (*vs.* 6), than children born just before these cutoff dates, *ceteris paribus*. This requirement is expected to hold, because—as discussed in Section 3—the January 1 and June 1 cutoffs greatly influence children's educational paths/opportunities in Hungary: first, the local DAB's evaluation or support is not needed for redshirting a child born on January 1, but the administrative requirement of the local DAB's evaluation is present for a child born on December 31; second, a child born on June 1 *vs.* May 31 is prescribed to enter primary school at age 7 *vs.* 6. Figure 1 foreshadows that the first-stage relationship is very strong and Appendix B formally tests it.

A sufficient condition for the *independence* assumption to hold is that birth month around the two cutoff points is effectively random. This requirement could fail, among others, in the following two notable cases. First, if children born on one side of the cutoff dates have above-average innate abilities than children born on the other side (i.e., if innate ability is not uniformly distributed across the birth months near the cutoffs). Second, it fails if parents with particular preferences tend to have June babies rather than May babies (i.e., tend to conceive their baby in September rather than August), and thus their child is likely to be among the older students in class. On the distribution of abilities and parental preferences, existing evidence is inconclusive;[26] but, given that I compare children born in two consecutive quarters, with one of the cutoffs being in the winter and the other in the summer season, differences between summer- and winter-born children, suggested by Bound and Jaeger (2001) and Buckles and Hungerman (2013), are not a concern. In Appendix C, I still test the relationship between being born after the cutoffs and observable characteristics, and check for bunching.

Regarding the *exclusion restriction*, it fails if children born after the appropriate cutoffs have systematically different outcomes for reasons other than an increased likelihood of starting school at age 7. A potential threat is that older students, all else equal (including SSA), tend to have higher test scores. As an example, consider the cutoff of January 1 and two otherwise identical children: Dora born in December and Julie born in January (of the next year). Suppose that both children are redshirted. Then, even though Julie had a higher *ex-ante* likelihood of being redshirted, Dora is 1 month older than Julie at the time of the test.[27] Thus, children born after January 1, such as Julie, may have different test scores for at least two reasons: a higher *ex-ante* likelihood of spending an extra year in childcare, but also being younger at the time of the test (*ceteris paribus*, including SSA).

The above concern can be eliminated by comparing two children born just before or just after the cutoff, by shrinking the window as much as possible, and selecting two otherwise identical children who were born just one day apart. Such an approach is possible when the exact day of birth is known, information that is unavailable in my data. At the same time, it is necessary to control for month of birth, to account for the steadily increasing share of children starting school at age 7, across the birth months (as in Figure 1), and to identify the impact *exclusively from the jump* in the share of redshirted children at the cutoffs. Thus, within a birth cohort I include children born in a 3-month window around the cutoffs, and control for trends in birth month linearly, allowing for different slopes on each side of the cutoffs. A linear estimator will under-(over-)estimate effects based on a true convex

[26] Dickert-Conlin and Elder (2010) tests whether parents systematically plan childbirth to capture the option value of sending their children to school at a younger age, thus avoiding an additional year of childcare costs, and they find no evidence of the option value influencing timing of birth. Also, using exact birthdates, McEwan and Shapiro (2008) and, despite a somewhat different context, McCrary and Royer (2011), also find no evidence of birth frequencies or students' observed socio-economic characteristics varying sharply around enrollment cutoff dates, suggesting that unobserved characteristics also vary smoothly around these dates. There also exists evidence suggesting that *quarter* of birth may be related to maternal and child characteristics; e.g., (i) Bound and Jaeger (2001) reviews potential factors associated with month of birth, such as health or personality traits, and find schizophrenia, intellectual disability, autism, and manic depression to vary across quarter of birth; (ii) Buckles and Hungerman (2013) documents that women giving birth in the winter months are younger, less educated, and less likely to be married. Huang et al. (2020) finds that more educated Chinese mothers tend to move their delivery date forward to have their children start school younger.

[27] Similarly, consider the June 1 cutoff and two otherwise identical children, the first born in March, the second in July. Suppose that the first child is redshirted and the second complies to the prescribed school-starting age of 7. Even though the second child had a higher likelihood of starting school at age 7, spending an additional, 4^{th} year in childcare, the first one is still 4.5 months older than the second at the time the test is taken.

(concave) first-stage relationship from the left and will over-(under-)estimate effects based on a true convex (concave) reduced-form from the left. In Appendix F I check robustness to a quadratic trend.

Exclusion would be also violated if there are other policies that use the same birthday cutoff, generating differential human capital investments on either side. As mentioned in Section 3, in my setting, besides SSA, the cutoffs also affects childcare-starting age and thus years spent in childcare, although to a much smaller extent—thus, Appendix H checks to what extent my main redshirting estimates are robust to accounting for the potentially endogenous years spent in childcare. It may also happen that because redshirted were considered not to be mature enough to start school on time, this judgment affects the behaviors of teachers, parents, and children; e.g., school administrators, teachers and parents may try to compensate redshirted children by providing additional resources like remedial tutoring or help with homework—thus, Appendix C checks and finds no evidence for such remediation.

I make the following four final remarks on identification and inference: (i) given that it is important to control for trends in birth month as discussed below, a 1-month window is not sufficient for identification; (ii) Appendix F checks to what extent my main results for the LATE of redshirting are robust to (a) a 1-month window around the January 1 cutoff, without a trend; (b) a 2-month window with a linear trend; and (c) a 3-month window with a quadratic trend around the cutoff; (iii) I do not use September 1 for identification, because on the left-side of that cutoff is SSA of 7 by default, and redshirting is not possible; (iv) based on Lee and Card (2008), the ideal level of clustering of standard errors (SEs) would be at the level of birthdate in my setup, but that would lead to an insufficient number of clusters (6). To avoid a downward bias in the SE estimates due to a small number of clusters, and as I expect the regressors and errors to be correlated across students within schools, I cluster SEs at the school level based on Cameron and Miller (2015). Appendix G checks and demonstrates that clustering at the school level is the most conservative relative to alternatives (clustering at birth month and cohort level) and accounting for the discrete running variable (Kolesar and Rothe, 2018).

5.4 Monotonicity, LATE, and Compliers with Their Average Characteristics

For the IV estimator to produce the LATE, the *monotonicity* requirement also has to hold. Monotonicity requires that, while the instrument may not influence school-starting age for some children, all of those who are influenced must be influenced in the same direction. Around the January 1 cutoff, monotonicity requires the following to hold for all children: if a child born before January 1 is redshirted (in the presence of the administrative barrier to redshirting), they would definitely be redshirted if they were born just a day later, on or after January 1 (in absence of this barrier). Since redshirting under the administrative barrier is more costly, this assumption is a non-controversial one. Around the June 1 cutoff, monotonicity requires, for all children, that if a child born before June 1 starts school at age 7 (equivalent to being redshirted), they would definitely start school at age 7 if born on or after June 1 (as required). Or, equivalently, it cannot happen that a child born before June 1 delays school entry from age 6 to 7, but is brought forward from 7 to 6 when being born after June 1.[28]

[28] Aliprantis (2012) and Barua and Lang (2016) argue that due to redshirting parents, monotonicity around the enrollment cutoff date could be violated. Given that this paper's focus is on redshirting using January 1 for identification, for which the monotonicity assumption is not controversial, and that children born after June 1 in the Hungarian setting

Given *relevance, independence, exclusion,* and *monotonicity,* the Wald estimand has the interpretation of the (local average) effect of starting school at age 7 (*vs.* 6), for the *compliers*, whose treatment status is influenced by the change in the instrument's value around the discontinuity cutoffs.

Around the January 1 cutoff, the compliers are *non-school-ready* children who are deterred from redshirting by the administrative barrier: these children would be redshirted if the parents were only required to discuss the issue of redshirting with the childcare teachers, but would not be redshirted if the parents had to ask for the local DAB's evaluation. For such parents, it is too costly to search for, travel to, and let the child take the school-readiness examination, either because the time costs are too high for them, or such parents are less informed and/or distrust the educational system. Potentially, there are also a group of complier children, whose parents *think* they are *non-school-ready* and who would be redshirted if only the childcare teachers' approval was needed, but who would not be advised to be redshirted by DAB developmental psychologists conducting the school-readiness evaluation.

Around the school enrollment cutoff of June 1, the compliers are *school-ready* children who enter primary school as prescribed by the regulation: at age 6 if born before June 1 (i.e., not redshirted), and at age 7 if born on or after June 1. That is, complier children around June 1 who start school a year older at age 7 do so solely because they were born after the enrollment cutoff and, had they been born before, would not have been redshirted and started school at a relatively young age of 6. (There might also be parents who never want to redshirt, even if their child was born before June 1 (say, in May) and was not school-ready. But, given in Hungary an overwhelming fraction (74 and 60 percent) of May-born boys and girls are redshirted (Figure 1), the share of such parents must be negligible.)

While it is impossible to identify the compliers in the data set—as they, and every other child, are observed in only the actual state of the world, and not in the counterfactual state—it is possible to assess their average characteristics. Following Almond and Doyle (2011), let us first define two independent binary variables, the instrument Z_j and the endogenous variable D^j, as follows:

$$Z_j = \left\{ \begin{array}{ll} 1 & \text{if born on/after the cutoff day} \\ 0 & \text{if born before the cutoff day} \end{array} \right\}; \; D^j = \left\{ \begin{array}{ll} 1 & \text{if starting school at the age of 7} \\ 0 & \text{if starting school at the age of 6} \end{array} \right\};$$

with j=1,2 corresponding to the January 1 and June 1 cutoffs, respectively. In addition, let us denote $D^j_{Z_j}$ as the value that D^j would take if Z_j were either 1 or 0. Then, compliers are those with $D^j_1 = 1$ and $D^j_0 = 0$. Let us denote their share π_{C_j}. Then, their observable characteristics can be written as

$$E\left(X|D^j_1 = 1, D^j_0 = 0\right) = \frac{\pi_{C_j} + \pi_{A_j}}{\pi_{C_j}} \left[E\left(X|D^j = 1, Z_j = 1\right) - \frac{\pi_{A_j}}{\pi_{C_j} + \pi_{A_j}} E\left(X|D^j = 1, Z_j = 0\right) \right], \tag{3}$$

where π_{A_j} is the share of always-takers (for whom $D^j_1 = 1$ and $D^j_0 = 1$) and $\pi_{C_j} = 1 - \pi_{A_j} - \pi_{N_j}$, where π_{N_j} is the share of never-takers (for whom $D^j_1 = 0$ and $D^j_0 = 0$). Defiers (for whom $D^j_1 = 0$ and $D^j_0 = 1$) are assumed away by the monotonicity assumption. Using independence between Z_j and D^j, the shares π_{A_j} (π_{N_j}) can be estimated by the empirical counterpart of $Prob(D^j_0 = 1)$ ($Prob(D^j_1 = 0)$).

are required to start school at age 7 but cannot be delayed further, these concerns are less relevant here.

5.5 Relationship to Existing Literature and Interpretation of LATE Estimates

This paper's identification strategy is most similar to the one used in Puhani and Weber (2007) and Altwicker-Hamori and Kollo (2012), who use predicted school entry age as an IV for the endogenous primary SSA, that is a function of the child's birth month, the primary school start month, and the enrollment cutoff month. Note that Puhani and Weber (2007)'s IV is of the same form as mine around the cutoff date of June 1.[29] Compared to them, I additionally control for the trend in birth month, allowing for different slopes on the two side of the discontinuity cutoffs, to account for different trends.

After establishing that this paper measures the causal impact of academic redshirting, for *non-school-ready* children, it is important to be transparent of what this paper does *not* measure. For this, I note that as I look at tests taken at a given stage of students' school career (at a given grade), the following collinearity holds (in months): $SSA + months\ since\ primary\ school\ entry = age\text{-}at\text{-}the\text{-}test$.

First, I do not attempt to disentangle the effect of SSA from the effect of age-at-the-test – the Hungarian institutional setup is not suitable for that as children are tested on the same date in a given grade. Black et al. (2011) disentangles these by using scores from Norwegian IQ tests taken outside of school and exploiting the variation in the mapping between birthdate and the year-of-the-test. I, however, look at the effect of SSA on grade repetition (retention) in further grades, as an outcome.

Second, I do not attempt to disentangle age-at-the-test effects from the impacts of time spent in school. There are institutional setups in other countries that provide a suitable natural experiment to compare children in the same grade with the same age at the time of the test, but who have spent different amounts of time in pre-school; e.g., Dustmann and Cornelissen (2019) exploits the variation in age-at-school-entry and the effective length of the first year in school in the UK, and Carlsson et al. (2015) exploits conditionally random variation in the assigned test date for IQ tests of 18-year-old Swedish men in preparation for military service, so that both age-at-the-test and number of days spent in school vary randomly after holding date of birth, parish, and expected graduation date fixed.[30]

Third, I do not compare a redshirted child's test scores with that of a non-redshirted child of the same age-at-the-test, because among non-grade-repeaters, there is no variation in SSA by age-at-the-test; the small variation in SSA by age-at-the-test that exists comes solely from grade-repeaters, is thus endogenous – I do not control for grade repetition, but I look at it as an outcome variable itself. Hence, my main redshirting estimate picks up the *combined* effect of starting school later – with one more year to mature before school entry – and being older at the time of the test.

In sum, I measure the effect of higher SSA—at age 7 *vs.* 6—, which may encompass both absolute (maturity) and relative age effects in class, and more years in childcare. But, I will show that my redshirting estimates are robust to accounting for time in childcare (Appendix H), are driven by absolute and not relative age effects (Appendix I), and there is no indication that teachers and parents compensate redshirted children with additional resources (e.g., remedial tutoring/help; Appendix C).

[29] Puhani and Weber (2007) does not control for a linear trend in month of birth—they only note that if b_i and s_i are exogenous, the theoretical SSA is exogenous and can be used as an instrument for the actual age of school entry (Puhani and Weber, 2007, at 367). They discuss factors that would invalidate their instrument, such as parental planning of conception, and acknowledge that month of birth "might exert some direct effect on physical and psychological health".

[30] The effect of longer exposure to kindergarten, without controlling for age-at-the-test, is measured by DeCicca and Smith (2013), which exploits the introduction of the "dual entry" scheme into kindergarten in 1990/1991 in BC, Canada.

6 Descriptive Statistics

6.1 Support for IV Validity

Figure 1 shows the raw first-stage relationship. Between children born in December *vs.* January and May *vs.* June, there is a large jump in the share of children with SSA of 7, of 17 and 28 percentage points, respectively; this share is then steadily increasing across birth months September–August.

Reassuring evidence supporting a smooth distribution of births around the January 1 cutoff can be seen in Appendix C. Table C2 shows no bunching on either side of the cutoff, and Table C3 shows the estimated coefficient on the discontinuity dummy $1\{X \geq x_d\}$ from model (1) on gender, parental education, family background variables listed in Section 4.2, and early childhood shocks presented in Table 2. None of the estimated $\hat{\beta}_1$'s differ from 0, providing reassuring support for the *independence* identification assumption and IV validity. That is, neither the fraction of boys, nor the share of children with low parental education or any other background characteristics, or with various developmental obstacles in early childhood, are discontinuously changing at the January 1 cutoff.

Before turning to the compliers and LATE estimates of higher SSA around the two cutoffs, I emphasize that based on Table C4, there is no indication that teachers and parents may try to compensate redshirted children by providing additional resources like remedial tutoring or help with homework; thus, the LATE estimates presented below will represent effects of higher SSA (either due to redshirting or due to the school enrollment cutoff), as opposed to effects of different treatments.

6.2 Characteristics of Redshirted Children and the Compliers

In this subsection, I show new evidence on factors associated with being redshirted, and who the compliers are around the January 1 and June 1 cutoffs for whom the LATEs are estimated.

Table 2 shows the shares of boys and girls born between October and March with different early childhood events and developmental obstacles, by whether having been redshirted and starting school at age 7 ($D = 1$) or at age 6 ($D = 0$). This table reveals four key relationships between adverse events in early childhood and D: (i) boys and girls are significantly more likely to be redshirted if they experienced a family shock (e.g., they were separated from their father before age 5, or the family had a negative income shock before then, or their parents rarely read to them), while (ii) girls are significantly more likely to be redshirted if they were separated from their mother before age 5. Furthermore, (iii) boys and girls are significantly more likely to be redshirted if they experienced a shock at birth (e.g., low birth weight, early-term delivery before week 36, or birth with complications), chronic illnesses in early childhood (e.g., sinusitis at ages 1-3), or said their first words only when they were 1 year old, or older. Finally, (iv) developmental problems at ages 4-5 (e.g., attention/cognition/verbal disorders, coordination problems) or diagnosis of ADHD are also significantly associated with the propensity of being redshirted at age 6 (for boys and attention disorders, significantly more associated). The length of having been breastfed is negatively related to being redshirted, yet only significantly at 10 percent.

Table 2 also reveals important gender differences in boys' and girls' developmental trajectories. The gender gap in speaking ability already opens up at the age of 1 (boys are more likely to say their first words at age 1 or older, and tend to say their first full sentence even later), and persists to age

4–5 (boys are more likely to have problems with speaking and cognition and are more likely to be diagnosed with ADHD). Boys are more likely to suffer from chronic ear illnesses before turning age 3.

Table C2 shows that on both sides of the January 1 cutoff, the share of children whose parents have at most primary or vocational education is higher for redshirted than on-time children (24 *vs.* 9 and 35 *vs.* 29 percents, respectively, for fall-borns). Thus, redshirting is more prevalent among low-status students—but, this relationship vanishes if early childhood indicators in Table 2 are controlled for.

Table 3 shows the average characteristics of the compliers. Around January 1, boys are over-represented among the compliers, but their other average characteristics are indistinguishable from the sample (except for number of siblings and whether they live in the capital). Around June 1, boys are under-, and those from better family backgrounds are over-represented, among the compliers.

Column (7) in *Panel B* of Table 3 shows the results of the direct test of whether compliers around the January 1 cutoff (for redshirting) are indeed less school-ready than compliers around the June 1 cutoff (the school enrollment cutoff date).[31] Compliers' share around the January 1 cutoff are significantly higher, precisely in the dimensions that have been found to be systematically related to redshirting and indicative of non-school-readiness: they are significantly more likely to have (i) been born with low birth weight, (ii) been separated from their mothers before their 1^{st} birthday, (iii) been older when they started to speak, (iv) suffered chronic ear illnesses in their first 3 years of life, (v) had cognition/verbal/coordination problems and been diagnosed with an attention disorder at ages 4-5.

Table 4 shows the average characteristics by gender: within the compliers around January 1, boys started to speak at even older age than girls, and they are more likely to have been diagnosed by ADHD at ages 4–5. Thus, boys among the compliers around January 1 are even more negatively selected than girls, and show more severe signs of non-school-readiness.

In sum, redshirted children are more likely to have suffered adverse developmental events in early childhood, and were less school-ready at age 6, on average, indicated by their lag in developmental markers at age 6. Also, compliers around the January 1 cutoff—relevant for academic redshirting—seem to have been indeed significantly less school-ready than compliers around the school enrollment cutoff of June 1, with complier boys seeming even less school-ready than complier girls.

7 LATE Estimates

7.1 LATE Estimates of Higher SSA due to Academic Redshirting

7.1.1 First-Stage Results

The left panel of Table B1 shows the first-stage results on starting school a year older due to redshirting, using the administrative test score data. 10^{th}-grader children born after January 1 are 11.2 percentage points more likely to start school at age 7, as opposed to 6, compared to children born before the cutoff (the baseline is 9.3 percent). The first-stage relationship is very strong, with no concerns of passing tests of weak instruments (Stock et al., 2002; Olea and Pflueger, 2013; Lee et al., 2022).

The first-stage relationship is significantly stronger for boys in absolute terms (with p-value $p =$

[31] The estimated share of compliers around January 1 is 32.5 percent, and the share of always-takers is 5.9 percent. The estimated share of compliers around June 1 is 53.5 percent, and the share of always-takers is 38.5 percent.

0.000), but not in relative terms: a boy (girl) in Grade 10 is 13 (9.5) percentage points more likely to be redshirted if born on or after January 1 than if born just before. However, given boys' higher baseline propensity to be redshirted relative to girls of around 4 percentage points (among those born in the fall), the effect for boys is around 115 percent, while for girls, it is around 125 percent.

7.1.2 Main LATE Estimates for Student Achievement in Grade 10, Secondary School Track, Educational Aspirations and Secondary School Graduation

The LATE estimates on student achievement in Grade 10 for all children are in Table 5. The first column in the left portion of *Panel A* reveals that redshirted children achieve a significantly higher score in mathematics, on average, by 0.12 SDs on Grade 10 testings, compared to on-time children (while the point estimate for reading test score is 0.09 SDs and on the propensity of repeating a grade is -3.24 percentage points, and they are not significantly different from zero).

The positive redshirting effect on student achievement in Grade 10 is only present for boys (*Panel B*): redshirted boys achieve a significantly higher score in mathematics and reading, on average, by 0.2 and 0.25 SDs, on Grade 10 testings, respectively, and they are 9.4 percentage points less likely than on-time children to repeat a grade by Grade 10 (which is a large effect size, compared to the control complier share of 11 percent). In contrast, the effect of redshirting for girls is insignificant for all outcomes. Formally testing the equality of boys' and girls' coefficient estimates gives $p = 0.019$, $p = 0.001$ and $p = 0.000$ for mathematics score, reading score and grade repetition, respectively.

Panel A of Table 6 show the impact of academic redshirting on secondary school track choice for all children. This crucial intermediate school outcome is measured by a binary variable indicating whether the student attends the j-th track at Grade 10; track #1 is the lowest track, granting a vocational degree but no high school diploma upon completion, without the option of continuing onto tertiary-level education; track #2 is the middle track, granting both a vocational and high school degree; track #3 is the highest (academic) track, granting a high school degree. Redshirting significantly decreases the propensity of going to the vocational track by 7.3 percentage points (49 percent), and significantly increases the propensity of going to the academic track by 5.6 percentage points (14 percent).

Again, the impact of redshirting on secondary school track choice is only significant for boys (*Panel B*): redshirted boys are significantly more likely to attend the academic track (by 11 percentage points, or 30 percent), and significantly less likely to attend the vocational track (by 11.9 percentage points, or 57 percent), while there are no detectable effects for girls. Formally testing the equality of boys' and girls' coefficient estimates gives $p = 0.001$ and $p = 0.005$ for the outcomes of attending the vocational and academic tracks, respectively. Given the 15 percentage points difference between boys' and girls' control complier share in the academic track, the estimated effect for boys implies that redshirting shrinks the baseline gender gap in attending the academic track by more than two-thirds.

Panel C of Table 6 displays the impact of redshirting on aspirations for highest level of educational attainment (i.e., less than high school, high school diploma, and more than high school by obtaining a tertiary degree) for all children. Academic redshirting significantly increases the propensity to aspire for a tertiary degree by 6.2 percentage points (11 percent of the baseline, control complier, mean).

Again, the positive redshirting impact on aspirations is only present for boys (*Panel D*): redshirted

boys are significantly less likely to aspire for less than a high school degree (by 5.2 percentage points, or 43 percent), significantly less likely to aspire for a high school degree (by 8.2 percentage points, or 22 percent), and significantly more likely to aspire for a tertiary degree (by 13.4 percentage points, or 26 percent), while there is no detectable effect for girls for any of the outcomes. Formally testing the equality of boys' and girls' coefficient estimates gives $p = 0.01$, $p = 0.002$ and $p = 0.009$ for the outcomes of aspiring for less than a high school degree, a high school degree, and a tertiary degree, respectively. Given the 6, 7, 13 percentage points difference between boys' and girls' baseline, control complier, shares in aspirations for $< HS$, HS, $HS+$, respectively, the estimated effects imply that redshirting could fully eliminate the gender gap in aspirations for the highest educational attainment.

Panel A of Table 7 displays the impact of academic redshirting on obtaining a vocational or high school degree for all: redshirting significantly increases the propensity to obtain a high school diploma by 16 percentage points (or 20 percent). Yet, again, the impact is only significant for boys (*Panel B*): redshirted boys are significantly more likely to obtain it by 19 percentage points (or 25 percent), while there is no significant effect for girls (with $p = 0.062$ for the difference in estimates), leading to academic redshirting more than closing the gender gap in obtaining a high school degree. (There is a negative effect on graduating with a vocational degree given redshirted children are less likely to end up in the vocational track to begin with, and overall there is no effect on dropping out of high school.)

7.1.3 Placebo and Robustness Checks

Appendix Table E1 shows the 1^{st}-stage and reduced-form estimates on the discontinuity dummy $Z = 1\{X \geq x_d\}$ (with D and various outcomes, respectively), using a 1- or a 2-month window, whenever possible, around placebo cutoffs (x_d : October 1,...,April 1). There are two main reassuring conclusions: (i) the biggest jump in the share of redshirted children is at the January 1 cutoff (an increase of 16.43 percentage points for boys); and (ii) the reduced-form estimates are either insignificant or have a different sign at the placebo cutoffs than the reduced-form estimate at the January 1 cutoff.

Then, Appendix Tables F1–F2 show that the point estimates and their precision do not change meaningfully when switching to a 2-month window (from a 3-month window) around the January 1 cutoff date, and when controlling for a quadratic trend in birth month with the 3-month window (*vs.* a linear trend). When restricting to a 1-month window, the point estimates are somewhat smaller, but the main conclusions still hold. Appendix Table G1 shows that my chosen level of clustering (at the school level) leads to larger, more conservative, standard errors and lower values of t-statistics than by clustering at the birth-month level or birth-month×year level, or accounting for the fact that there are only 3 months of birth of the two sides of the cutoffs (Kolesar and Rothe, 2018).

Appendix H shows that my main redshirting estimates are robust to accounting for *years in childcare*, a decision variable of parents which, besides SSA, might also change at the January 1 cutoff. Indeed, Table H1 shows that the share of children who start childcare at the age of 3 or later jumps by 19 percentage points at the January 1 cutoff,[32] and the average years spent in childcare is 0.092 *lower* among January-born children than among December-born ones. Thus, the cutoff affects to some

[32] 28.7 percent of December-born children start childcare at the age of 2, but only 9.3 percent of January-born do so. The share of children who start childcare at the age of 4 or later jumps by 4.5 percentage points at the January 1 cutoff.

extent *years in childcare*. Given the identity SSA =*years in childcare*+CSA, I jointly estimate the effects of the first two, where I use the fraction of 3-year-olds in childcare in the municipality when the child was 2 years old, as an additional source of exogenous variation, stemming from the institutional background.[33] Standard tests support IV strength and exogeneity (Tables H2 –H3).

The strong LATE estimates of redshirting for boys prevail even if accounting for *years in childcare*; e.g., even holding *years in childcare* fixed, redshirted boys achieve a significantly higher score in mathematics and reading, on average, by 0.26 and 0.28 SDs on Grade 10 testings, respectively, they are 7.4 percentage points less likely than on-time children to repeat a grade by Grade 10, are significantly more likely to attend the academic track (by 11 percentage points) and aspire for a tertiary degree (by 13 percentage points). For girls there are still no positive impacts for redshirting (Table H3).

Finally, Tables I1 and I2 in Appendix I show that my main estimates are more conservative than when also accounting for relative age (RA) effects in class, through which higher SSA may manifest. My unique data on all students' birthdate and class identifiers in a cohort allows me to construct a precise measure of each students' within-class RA rank, to separate absolute age effects from within-class RA effects. By exploiting natural variation in the share of summer-born children in the class as well as month of birth, I horse-race absolute and relative age effects behind starting school a year older due to redshirting. I find the positive redshirting effects on student achievement to be driven solely by absolute age effects (as RA effects are all insignificant). Thus, redshirting matters as it makes children older in absolute terms, rather than by only making them older relative to their classmates.

7.2 LATE Estimates of Higher SSA due to the School Enrollment Cutoff Date

For mathematics and reading test scores, the estimates are 0.14 and 0.18 SDs in Grade 10, respectively (see the right portion of *Panel A* in Table 5)—these estimates are consistent with existing literature, e.g., with Puhani and Weber (2007), Altwicker-Hamori and Kollo (2012), McEwan and Shapiro (2008), and Elder and Lubotsky (2009). Also, children who start school at the age of 7 just because they were born after and comply with the enrollment cutoff date of June 1 are less likely to fail a grade (by 5.8 percentage points by Grade 10, or by 51 percent), aspire for more education (are 4 percentage points less likely to aspire for at most a high school degree, and are 8 percentage points more likely to aspire for a tertiary degree; see *Panel C* in Table 6), and are more (less) likely to enrol in academic (vocational) high school tracks (by 6 and 8 percentage points, respectively, see *Panel A* in Table 6), and are 5 percentage points more likely to obtain a high school degree (see *Panel A* in Table 7). Notably, the significantly positive effects of higher SSA due to complying with the school enrollment cutoff date of June 1 typically disappears when accounting for years spent in childcare (see Table H3).

In sum, the LATE estimates of higher SSA due to redshirting (of Section 7.1) are, on average, very similar to that of due to the school enrollment cutoff (of Section 7.2). But, there is a key policy-relevant

[33]The order in which parents make decisions about their children's educational path is ambiguous: they may decide sequentially, first for CSA, and only then consider redshirting at the relevant SSA; or, they may very well be forward-looking, and decide on CSA with the possibility of redshirting in mind. With the former, one would want to control for CSA, as in Table H3. With the latter, CSA should not be a control, as the shorter time spent in childcare is part of the mechanism through which redshirting impacts child outcomes. Given the ambiguity regarding the timing of parental decisions, I present the regression that accounts for *years spent in childcare* as a robustness check only.

difference on which groups of children are driving the positive impacts of higher SSA: redshirting has significantly positive effects only for *boys* (and significantly larger than for girls), but, both boys and girls experience significantly positive effects of higher SSA due to the enrollment cutoff.[34]

8 Mechanisms

8.1 Student Achievement in Primary School

Academic redshirting has a significantly positive impact on test scores already in primary school. *Panel A* of Table D1 reveals that redshirted children achieve a significantly higher score in mathematics, on average, by 0.21 and 0.145 SDs, on Grade 6 and 8 testings, respectively, compared to on-time children. The impacts on the Grade 6 and 8 reading scores are 0.2 and 0.13 SDs, respectively (*Panel C*).

The positive redshirting effect on test scores is only present for boys; e.g., boys' effect of redshirting is 0.3 and 0.14 SDs, on the Grade 6 and 8 mathematics testings, respectively, compared to on-time children (*Panel B*). In contrast, the effect of redshirting for girls is only significant at the 10 percent level on their Grade 8 mathematics score. For reading, the corresponding effects are 0.29 and 0.22 SDs, respectively, for boys, and all point estimates are small and insignificant for girls (*Panel D*).

Table D2 reveals that redshirted children are 3 percentage points less likely than on-time children to repeat a grade by Grade 6. The effect size is large compared to the baseline grade repeaters' share of 3.6 percent, even though there are no detectable effects by Grade 8. Again, the impact of redshirting on grade repetition is only significant for boys; it is -5.1 percentage points by Grade 6; for girls, all point estimates are much smaller (in absolute terms) and none of them are statistically significant.[35]

In sum, redshirting helps boys perform better in secondary school, aspire for higher education levels and graduate from high school *because* it helps them perform better already in primary school.

8.2 Mental Health (Anxiety), Confidence, and Student Well-Being

Having shown that the positive impacts of redshirting on student outcomes are significant only for boys, I now turn to mental health outcomes, to shed further light on the underlying mechanisms.

Academic redshirting has a significantly positive impact on reducing boys' symptoms of anxiety. The left of *Panel A* of Table 8 reveals that redshirted boys have lower anxiety scores than on-time children by over 1 SD in Grade 8, and are 46 percentage points less likely to experience frequent symptoms of Generalized Anxiety Disorder (GAD). Again, none of the corresponding estimates for girls are significantly different from zero, and testing the equality of boys' and girls' coefficient estimates gives $p = 0.054$ and $p = 0.11$ for the anxiety score and the likelihood of experiencing frequent symptoms of GAD, respectively. Column (3) conforms these results, using administrative health data: higher school-starting age due to academic redshirting significantly decreases expenditures on prescription

[34] For boys none of the effects in Tables 5–7 are different *across the cutoffs*, while for girls all effects are significantly higher at the June 1 than the January 1 cutoff at the 5 percent level in Tables 5–6, and they are larger at the 10 percent level in Table 7 for $1\{VOCdegr\}$ and $1\{HSdegr\}$, with $p = 0.091$ and $p = 0.087$, respectively.

[35] Formally testing the equality of boys' and girls' coefficient estimates gives $p = 0.01$ and $p = 0.52$ ($p = 0.052$ and $p = 0.002$) for mathematics (reading) score in Grade 6 and 8, respectively; and $p = 0.02$ and $p = 0.2$ for grade repetition.

psycholeptics—primarily used for treating anxiety—by 4 percent for boys, and this point estimate is significantly larger at the 5 percent level ($p = 0.042$) than the estimate of girls (which is insignificant).

Academic redshirting has a significantly positive impact on boys' confidence and a negative impact on being bullied in class. The left of *Panel B* of Table 8 reveals that redshirted boys' propensity to feel unvalued and untalented decreases, and they are significantly less likely to be bullied in class for their appearance or poor academic performance. Girls' estimates are all insignificant, and the respective p-values for testing the equality of estimates across genders are $p = 0.07$, $p = 0.2$, and $p = 0.04$. These results are also in line with findings of boys generally benefiting more from environments that improve children's socio-emotional skills (Walton and Cohen, 2011; Sisk et al., 2018).

These results suggest that mental health is an important mechanism behind why academic redshirting improves boys' student achievement through secondary school, and increases their educational aspirations and actual attainment. For boys, extra time before school seems to lead to better mental health (in terms of being less anxious), increased self-confidence (in terms of feeling valued and talented) and better overall well-being. Taking into account that boys fall behind already at ages 1–4 in their speaking, hyperactivity and attention abilities, and are less likely to be school-ready, academic redshirting also helps boys to overcome non-school-readiness, and thus leads to a boost in boys' human capital and increased maturity. The aforementioned factors—mental well-being, self-confidence, maturity and conquered developmental obstacles—are all associated with better student achievement.

9 Conclusion

This paper is the first to measure the causal impact of academic redshirting, using Hungarian administrative test score and medical data, and survey data on mental health. Exploiting an administrative barrier—a school-readiness evaluation—compulsory for potentially redshirted fall-born children, I identify a new LATE of higher SSA due to redshirting, for non-school-ready children, whom an extra year before school could help overcome developmental deficits, but who are deterred from redshirting by the barrier. In contrast, in the Hungarian setting in which children born before the school enrollment cutoff date could be redshirted, the LATE of higher SSA due to (complying with) the enrollment cutoff—the focus of existing literature—is for school-ready children, who start school a year older just because they were born after the cutoff, but who would have started on time if born just before.

Understanding redshirting brings a new and important angle on school-readiness, as it is relevant for parental decision-making and for policy. First, direct evidence on the impacts of academic redshirting may be more relevant for parents (who are considering redshirting their children, given their birthdate and skill levels around age 6) and policy-makers (who need to regulate parents' opportunities for academic redshirting), than the existing evidence on higher SSA due to (complying with) the enrollment cutoff. Second, whether these two types of LATEs for higher SSA—due to redshirting *vs.* the school enrollment cutoff—are being driven by the same group of children is an empirical question.

Besides providing a separate cutoff for identification of the causal effect of academic redshirting, the Hungarian setting is unique also since: (i) childcare is free, thus the counterfactual care arrangement is public care and there are no direct costs of redshirting for parents; (ii) redshirters are from neither

poor nor rich socio-economic backgrounds, when controlling for the fact that that they are more likely to have faced developmental obstacles in early childhood. Even though the Hungarian setting is different from typical US settings, it provides the opportunity to measure the effect of redshirting for non-school-ready children from any backgrounds, consistent with the original policy aim of redshirting.

I find that higher SSA due to academic redshirting (using the January 1 cutoff) leads to higher test scores in Grades 6-10 by 0.1-0.2 SD and higher educational aspirations, and a higher propensity to attend the academic high school track and to obtain a high school degree, but these positive redshirting effects are significant only for boys. I present new evidence that boys fall behind already at ages 1–4 in their speaking, hyperactivity and attention abilities, and are more likely to be non-school-ready than girls. In fact, non-school-ready boys are over-represented among the compliers, for whom the positive redshirting effects are estimated. I find that redshirted boys are less anxious, more confident and less prone to being bullied in class. Given that redshirting sets non-school-ready boys—but not girls—up to be college bound in tracks, and closes the gender gap in aspirations and in obtaining a high school degree, it has the potential to mitigate the "boys' crisis" in education, by helping boys overcome non-school-readiness and improve their student achievement and mental health.

In sum, boys gain from a later school start irrespective of whether it is due to redshirting barriers or school entry rules, but girls only gain if it is due to the school entry rules (nevertheless, girls do not seem to lose from being redshirted). Given that children around the January 1 cutoff, relevant for redshirting, are older than children around the school enrollment cutoff of June 1, it seems that young school-ready girls benefit from an extra year in their student achievement, but older and moderately non-school-ready girls do not. At the same time, both young school-ready boys and older non-school-ready boys seem to benefit from an extra year in their student achievement *and* mental health.

Future research may want to formally model parents' redshirting decision (as in, e.g., Ricks (2022) or Datar (2006b)), with an emphasis on their beliefs on their children's school-readiness and the returns to academic redshirting for them. One intriguing finding of this paper is that girls, on average, do not seem to benefit from redshirting, yet a significant proportion of them are still redshirted. Young school-ready girls' LATE estimates for higher SSA due to the school enrollment cutoff is significantly positive, suggesting that they do experience the "gift-of-age" gains. The two LATE estimates for girls, at first glance, suggest that either some *actually school-ready* girls are redshirted despite no need for it, or that parents may be unaware of/underestimate the potentially negative effects of holding back an older, to some extent non-school-ready girl, opening up interesting avenues for further research.

References

Abdulkadiroglu, A., P. A. Pathak, and C. R. Walters (2018). Free to Choose: Can School Choice Reduce Student Achievement? *American Economic Journal: Applied Economics 10*(1), 175â206.

Aliprantis, D. (2012). Redshirting, Compulsory Schooling Laws, and Educational Attainment. *Journal of Educational and Behavioral Statistics 37*(2), 316–338.

Aliprantis, D. (2014). When Should Children Start School? *Journal of Human Capital 8*(4), 481–536.

Almond, D. and J. J. Doyle (2011). After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays. *American Economic Journal: Economic Policy 3*(3), 1–34.

Altwicker-Hamori, S. and J. Kollo (2012). Whose Children Gain from Starting School Later? Evidence from Hungary. *Educational Research and Evaluation: An International Journal on Theory and Practice) 18*(5), 459–488.

Anderson, M. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.

Arnold, G. and B. Depew (2018). School-Starting Age and Long-Run Health in the United States. *Health Economics 27*(12), 1904–1920.

Attar, I. and D. Cohen-Zada (2018). The Effect of School Entrance Age on Educational Outcomes: Evidence Using Multiple Cutoff Dates and Exact Date of Birth. *Journal of Economic Behavior and Organization 153*, 38–57.

Aud, S., S. Wilkinson-Flicker, P. Kristapovich, A. Rathbun, X. Wang, and J. Zhang (2013). The Condition of Education 2013. Working Paper 37, National Center for Education Statistics.

Autor, D., D. Figlio, K. Karbownik, J. Roth, and M. Wasserman (2019). Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. *American Economic Journal: Applied Economics 11*(3), 338–381.

Balestra, S., B. Eugster, and H. Liebert (2020). Summer-Born Struggle: The Effect of School Starting Age on Health, Education, and Work. *Health Economics 29*(5), 591–607.

Ballatore, R. M., M. Paccagnella, and M. Tonello (2020). Bullied Because Younger than my Mates? The Effect of Age Rank on Victimisation at School. *Labour Economics 62*.

Barua, R. and K. Lang (2016). School Entry, Educational Attainment, and Quarter of Birth: A Cautionary Tale of a Local Average Treatment Effect. *Journal of Human Capital 10*(3), 347–376.

Bassok, D. and S. F. Reardon (2013). "Academic Redshirting" in Kindergarten Prevalence, Patterns, and Implications. *Educational Evaluation and Policy Analysis 35*(3), 283–297.

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

Bertrand, M. and J. Pan (2013). The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics 5*(1), 32–64.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *Review of Economics and Statistics 93*(2), 455–467.

Bound, J. and D. A. Jaeger (2001). Do Compulsory School Attendance Laws Alone Explain the Association Between Quarter of Birth and Earnings? *Research in Labor Economics 19*, 83–108.

Brandlistuen, R. E., M. Flato, C. Stoltenberg, S. S. Helland, and M. V. Wang (2021). Gender Gaps in Preschool age: A Study of Behavior, Neuro-Development and Pre-academic Skills. *Scandinavian Journal of Public Health 49*(5), 503–510.

Buckles, K. S. and D. M. Hungerman (2013). Season of Birth and Later Outcomes: Old Questions, New Answers. *The Review of Economics and Statistics 95*(3), 711–724.

Cameron, A. C. and D. L. Miller (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources 50*(2), 317–372.

Cannon, J. S. and S. Lipscomb (2008, May). Changing the Kindergarten Cutoff Date: Effects on California Students and Schools. Occasional papers, Public Policy Institute of California.

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

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

Cook, P. J. and S. Kang (2016). Birthdays, Schooling, and Crime: Regression-Discontinuity Analysis of School Performance, Delinquency, Dropout, and Crime Initiation. *American Economic Journal: Applied Economics 8*(1), 33–57.

Cook, P. J. and S. Kang (2018). The School-Entry-Age Rule Affects Redshirting Patterns and Resulting Disparities in Achievement. Working Paper 24492, National Bureau of Economic Research.

Cook, P. J. and S. Kang (2020). Girls to the Front: How Redshirting and Test-score Gaps Are Affected by a Change in the School-Entry Cut Date. *Economics of Education Review 76*, 1019–68.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schonberg (2017). Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance. *Journal of Political Economy 126*, 2356–2409.

Cornwell, C., D. Mustard, and J. Van Parys (2013). Noncognitive Skills and the Gender Disparities in Test Scores and Teacher Assessments: Evidence from Primary School. *Journal of Human Resources 48*(1), 236–64.

Crawford, C., L. Dearden, and C. Meghir (2010). When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England. IFS Working Paper 10/06.

Datar, A. (2006a). Does Delaying Kindergarten Entrance Give Children a Head Start? *Economics of Education Review 25*(1), 43–62.

Datar, A. (2006b). The Impact of Kindergarten Entrance Age Policies on the Childcare Needs of Families. *Journal of Policy Analysis and Management 25*(1), 129–153.

Datar, A. and M. A. Gottfried (2015). School Entry Age and Children's Social-Behavioral Skills: Evidence From a National Longitudinal Study of U.S. Kindergartners. *Educational Evaluation and Policy Analysis 37*(3), 333–353.

DeCicca, P. and J. Smith (2013). The Long-Run Impacts of Early Childhood Education: Evidence from a Failed Policy Experiment. *Economics of Education Review 36*(C), 41–59.

Dee, T. S. and H. H. Sievertsen (2015). The Gift of Time? School Starting Age and Mental Health. *Health Economics 27*(5), 781–802.

Deming, D. and S. Dynarski (2008). The Lengthening of Childhood. *Journal of Economic Perspectives 22*(3), 71–92.

Dhuey, E., D. Figlio, K. Karbownik, and J. Roth (2019). School Starting Age and Cognitive Development. *Journal of Policy Analysis and Management 38*(3), 538–578.

Dhuey, E. and S. Lipscomb (2008). What Makes a Leader? Relative Age and High School Leadership. *Economics of Education Review 27*(2), 173–183.

Dhuey, E. and S. Lipscomb (2010). Disabled or Young? Relative Age and Special Education Diagnoses in Schools. *Economics of Education Review 29*(5), 857–872.

Dickert-Conlin, S. and T. Elder (2010). Suburban Legend: School Cutoff Dates and the Timing of Births. *Economics of Education Review 29*(5), 826–841.

DiPrete, T. A. and J. L. Jennings (2012). Social and Behavioral Skills and the Gender Gap in Early Educational Achievement. *Social Science Research 41*(1), 1–15.

Dobkin, C. and F. Ferreira (2010). Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes? *Economics of Education Review 29*(1), 40–54.

Dong, Y. (2010). Kept Back to Get Ahead? Kindergarten Retention and Academic Performance. *European Economic Review 54*(2), 219–236.

Du, Q., H. Gao, and M. D. Levi (2012). The Relative-Age Effect and Career Success: Evidence from Corporate CEOs. *Economics Letters 117*(3), 660–662.

Dustmann, C. and T. Cornelissen (2019). Early School Exposure, Testscores, and Noncognitive Outcomes. *American Economic Journal: Economic Policy 11*(2), 35–63.

Elder, T. and D. Lubotsky (2009). Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers. *Journal of Human Resources 44*(1), 641–683.

Elder, T. E. (2010). The Importance of Relative Standards in ADHD Diagnoses: Evidence Based on Exact Birth Dates. *Journal of Health Economics 29*(5), 641–656.

Evans, W. N., M. S. Morrill, and S. T. Parente (2010, September). Measuring Inappropriate Medical Diagnosis and Treatment in Survey Data: The Case of ADHD Among School-age Children. *Journal of Health Economics 29*(5), 657–673.

Fertig, M. and J. Kluve (2005). The Effect of Age at School Entry on Educational Attainment in Germany. Working Paper, IZA DP No. 1507.

Fidjeland, A., M. Rege, I. F. Solli, and I. Storksen (2023). Reducing the Gender Gap in Early Learning: Evidence from a Field Experiment in Norwegian Preschools. *European Economic Review 154*, 104413.

Fortin, N., K. Chen, P. Oreopoulos, and S. Phipps (2013). Young in Class: Implications for Inattentive/Hyperactive Behaviour of Canadian Boys and Girls. *Canadian Journal of Economics 48*, 1601–1634.

Fortin, N. M., P. Oreopoulos, and S. Phipps (2015). Leaving Boys Behind: Gender Disparities in High Academic Achievement. *Journal of Human Resources 50*(3), 549–579.

Fredriksson, P. and B. Ockert (2014). Life-Cycle Effects of Age at School Start. *The Economic Journal 124*(579), 977–1004.

Fumarco, L., S. A. Hartmann, and F. Principe (2024). A Neglected Determinant of Eating Behaviors: Relative Age. Working Paper, IZA DP No. 16920.

Golsteyn, B. H. and T. Schils (2014). Gender Gaps in Primary School Achievement: A Decomposition Into Endowments and Returns to IQ and Non-cognitive Factors. *Economics of Education Review 41*, 176–187.

Goodman, A. and B. Sianesi (2005). Early Education and Children's Outcomes: How Long Do the Impacts Last? *Fiscal Studies 26*(4), 513–548.

Gorlitz, K., M. Penny, and M. Tamm (2022). The Long-term Effect of Age at School Entry on Cognitive Competencies in Adulthood. *Journal of Economic Behavior Organization 194*, 91–104.

Gortz, M., E. R. Johansen, and M. Simonsen (2018). Academic Achievement and the Gender Composition of Preschool Staff. *Labour Economics 55*, 241–258.

Graue, B. and J. DiPerna (2000). Redshirting and Early Retention: Who Gets the "Gift of Time" and What Are Its Outcomes? *American Educational Research Journal 37*(2), 509–534.

Haeck, C., R. Lacroix, C. Montmarquette, and R. E. Tremblay (2023). *La Sous-Scolarisation des Hommes et le Choix de Profession des Femmes*. Les Presses de l'Universite de Montreal.

Havnes, T. and M. Mogstad (2015). Is Universal Child Care Leveling the Playing Field? *Journal of Public Economics 127*(C), 100–114.

Hemelt, S. W. and R. B. Rosen (2016). School Entry, Compulsory Schooling, and Human Capital Accumulation: Evidence from Michigan. *The B.E. Journal of Economic Analysis Policy 16*(4), 2015–219.

Horvath, H. (2015). Classroom Assignment Policies and Implications for Teacher Value-Added Estimation. Working Paper, University College London, Institute of Education.

Huang, C., S. Zhang, and Q. Zhao (2020). The Early Bird Catches the Worm? School Entry Cutoff and the Timing of Births. *Journal of Development Economics 143*(C).

Jenkins, J. M. and C. K. Fortner (2017). Kindergarten Redshirting: Motivations and Spillovers Using Census-Level Data. *Early Childhood Research Quarterly 108*(38), 44–56.

Jenkins, J. M. and C. K. Fortner (2019). Forced to Redshirt: Quasi-Experimental Impacts of Delayed Kindergarten Entry. Technical report, EdWP 19-120. https://doi.org/10.26300/mb5s-p142.

Johansen, E. R. (2021). Relative Age for Grade and Adolescent Risky Health Behavior. *Journal of Health Economics 76*.

Kolesar, M. and C. Rothe (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review 108*(8), 2277–2304.

Landerso, R., H. S. Nielsen, and M. Simonsen (2017). School Starting Age and the Crime-Age Profile. *The Economic Journal 127*(602), 1096–1118.

Lee, D. S. and D. Card (2008). Regression Discontinuity Inference with Specification Error. *Journal of Econometrics 142*(2), 655–674.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022). Valid t-Ratio Inference for IV. *American Economic Review 112*(10), 3260–90.

Lenard, M. and P. Pena (2018). Maturity and Minorities: The Impact of Redshirting on Achievement Gaps. *Education Economics 26*, 1–17.

Lubotsky, D. and R. Kaestner (2016). Do 'Skills Beget Skills'? Evidence on the Effect of Kindergarten Entrance Age on the Evolution of Cognitive and Non-Cognitive Skill Gaps in Childhood. *Economics of Education Review 53*, 194–206.

Magnuson, K. and G. J. Duncan (2016). Can Early Childhood Interventions Decrease Inequality of Economic Opportunity? *RSF: The Russell Sage Foundation Journal of the Social Sciences 2*(2), 123–141.

McAdams, J. M. (2016). The Effect of School Starting Age Policy on Crime: Evidence from U.S. Microdata. *Economics of Education Review 54*(579), 227–241.

McCrary, J. and H. Royer (2011). The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth. *American Economic Review 101*(1), 158–95.

McEwan, P. J. and J. S. Shapiro (2008). The Benefits of Delayed Primary School Enrollment: Discontinuity Estimates Using Exact Birth Dates. *Journal of Human Resources 43*(1), 1–29.

Muhlenweg, A., D. Blomeyer, H. Stichnoth, and M. Laucht (2012). Effects of Age at School Entry (ASE) on the Development of Non-cognitive Skills: Evidence from Psychometric Data. *Economics of Education Review 31*(3), 68–76.

Muhlenweg, A. M. and P. A. Puhani (2010). The Evolution of the School-Entry Age Effect in a School Tracking System. *Journal of Human Resources 45*(2), 407–438.

Murphy, R. and F. Weinhardt (2020). Top of the Class: The Importance of Ordinal Rank. *The Review of Economic Studies 87*(6), 2777–2826.

Olea, J. L. M. and C. Pflueger (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Oosterbeek, H., S. ter Meulen, and B. van der Klaauw (2021). Long-Term Effects of School-Starting-Age Rules. *Economics of Education Review 84*, 102144.

Patalay, P., J. Belsky, P. Fonagy, P. Vostanis, N. Humphrey, J. Deighton, and M. Wolpert (2015). The Extent and Specificity of Relative Age Effects on Mental Health and Functioning in Early Adolescence. *Journal of Adolescent Health 57*(5), 475–481.

Pena, P. A. (2017). Creating Winners and Losers: Date of Birth, Relative Age in School, and Outcomes in Childhood and Adulthood. *Economics of Education Review 56*, 152–176.

Pena, P. A. and A. L. Duckworth (2018). The Effects of Relative and Absolute Age in the Measurement of Grit from 9th to 12th Grade. *Economics of Education Review 66*(C), 183–190.

Ponzo, M. and V. Scoppa (2014). The Long-Lasting Effects of School Entry Age: Evidence from Italian Students. *Journal of Policy Modeling 36*(3), 578–599.

Puhani, P. A. and A. Weber (2007). Does the Early Bird Catch the Worm? *Empirical Economics 32*(2-3), 359–386.

Reeves, R. V. (2022). Of Boys and Men: Why the Modern Male Is Struggling, Why It Matters, and What to Do About It. *Brookings Institution Press*.

Ricks, M. D. (2022). Strategic Selection Around Policy Recommendations: The Case of Kindergarten Entry. Technical report, Retrieved from the Author's Website: https://www.michaeldavidricks.com/research.

Schanzenbach, D. W. and S. H. Larson (2017). Is Your Child Ready for Kindergarten? 'Redshirting' May Do More Harm. *Education Next 17*(3).

Schneeweis, N. and M. Zweimuller (2014). Early Tracking and the Misfortune of Being Young. *Scandinavian Journal of Economics 116*, 394–428.

Sisk, V. F., A. P. Burgoyne, J. Sun, J. L. Butler, and B. N. Macnamara (2018). To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? *Psychological Science 29*(4), 549–571.

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 B.E. Journal of Economic Analysis and Policy 9*(1).

Sprietsma, M. (2010). Effect of Relative Age in the First Grade of Primary School on Long-Term Scholastic Results: International Comparative Evidence Using PISA 2003. *Education Economics 18*(1), 1–32.

Stangeland, E. B., K. Lundetrae, and E. Reikeras (2018). Gender Differences in Toddlers' Language and Participation in Language Activities in Norwegian ECEC Institutions. *European Early Childhood Education Research Journal 26*(3), 375–392.

Stipek, D. (2002). At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents. *Social Policy Report 16*(2), 3–17.

Stock, J. H., J. H. Wright, and M. Yogo (2002). A Survey of Weak Instruments and Weak Identification in GMM. *Journal of Business and Economic Statistics 20*(4), 518–529.

Szabo-Morvai, A., D. Horn, A. Lovasz, and K. De Witte (2023). Universal Preschool and Cognitive Skills - The Role of School Starting Age as a Moderating Factor. *Early Childhood Research Quarterly 64*, 278–289.

Walton, G. M. and G. L. Cohen (2011). A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students. *Science 331*(3), 1447–1451.

Tables

Table 2: **Fraction of Children with Different Early Childhood and Developmental Obstacles Born Between October and March (Around the January 1 Cutoff), by Gender and School-Starting Age, Survey Data**

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: around birth	***premature birth***	***birth complications***	***born w/ low birth weight***	***breastfed <3 months***	***separated from mother***	***separated from father***
boy	0.0028	0.0226**	-0.0171**	0.0289*	0.0108**	0.0028
	[0.008]	[0.011]	[0.008]	[0.015]	[0.004]	[0.011]
D	**0.0644*** **	**0.0336***	**0.0919*** **	**0.0380***	**0.0297*** **	**0.0378** **
	[0.016]	[0.018]	[0.018]	[0.022]	[0.011]	[0.018]
boy × *D*	-0.0209	-0.0343	-0.0378*	-0.0405	-0.0337***	0.0005
	[0.020]	[0.024]	[0.022]	[0.030]	[0.012]	[0.024]
const	0.0649***	0.1285***	0.0817***	0.2761***	0.0147***	0.1281***
	[0.006]	[0.008]	[0.006]	[0.011]	[0.003]	[0.007]
Panel B: 1-3y old	1^{st} ***words > 1y old***	1^{st} ***sentences age in months***	***chronic illness in <3y face***	***chronic illness in <3y ear***	***few reading few books***	***income shock***
boy	0.0679***	1.6059***	0.0111	0.0210**	0.0122	0.0038
	[0.015]	[0.214]	[0.007]	[0.009]	[0.011]	[0.007]
D	**0.0574*** **	**1.4866*** **	**0.0027**	**0.0586*** **	**0.0604*** **	**0.0343** **
	[0.022]	[0.425]	[0.009]	[0.016]	[0.018]	[0.015]
boy × *D*	-0.0137	0.1677	0.0187	-0.0348*	-0.0352	-0.0182
	[0.030]	[0.554]	[0.014]	[0.021]	[0.023]	[0.019]
const	0.2651***	20.5471***	0.0381***	0.0907***	0.1011***	0.0695***
	[0.011]	[0.159]	[0.005]	[0.007]	[0.008]	[0.006]
Panel C: 4-5y old	***problems with nervsys***	***problems with speaking***	***problems with cognition***	***problems with coord***	***hyper-activity***	***attention disorder***
boy	0.0032	0.0628***	0.0083***	0.0015	0.0173***	0.0217***
	[0.002]	[0.012]	[0.003]	[0.002]	[0.003]	[0.004]
D	**0.0109** **	**0.0847*** **	**0.0152*** **	**0.0093***	**0.006**	**0.0133** **
	[0.005]	[0.019]	[0.005]	[0.005]	[0.004]	[0.006]
boy × *D*	-0.0103*	-0.0188	0.0019	-0.0017	0.0056	0.0162*
	[0.006]	[0.027]	[0.008]	[0.006]	[0.009]	[0.010]
const	0.0040***	0.1400***	0.0060***	0.0052***	0.0065***	0.0092***
	[0.001]	[0.008]	[0.001]	[0.001]	[0.002]	[0.002]

Notes: [1] This table shows the estimated coefficients and standard errors of a given outcome Y (seen in bold in the columns) on *boy*, $D = 1\{$***start school at 7***$\}$, and their interactions (without control variables).
[2] In *Panel A*, outcomes pertain to events/shocks around birth: (1) the child was born before 37 weeks of pregnancy have been completed; (2) birth complications occurred at birth; (3) child was born with very low birth weight /<2kg/; (4) the child was breastfed for less than 3 months; (5)-(6) the child was separated from her mother-father for at least a month in the first 1 year.
Panel B pertains to events/shocks/obstacles in the child's first 3 years of life: (1) the child said the first words being more than 1-year-old; (2) age in months when the child said the first complete sentences; (3)-(4) the child had a chronic face-ear illness; (5) only a few times per month the child was read to from books; (6) the family suffered a major income shock.
Panel C pertains to events/shocks/obstacles at ages 4-5: (1) child was diagnosed with problems with nervous system; (2) child was diagnosed with problems with speaking; (3) child was diagnosed with problems with cognition; (4) child was diagnosed with problems with movement and coordination; (5) child was diagnosed as hyperactive; (6) child was diagnosed with attention disorder.
[3] Standard errors are clustered at the school-level, *** p<0.01, ** p<0.05, * p<0.1.
[4] The sample includes children born between October and March (around the January 1 'redshirting' cutoff).
[5] The estimate for the *const* shows, for a given outcome, the fraction of non-redshirted girls, and the coefficient estimate for *boy* shows to what extent, on average, the share of non-redshirted boys differs from that of non-redshirted girls', while the coefficient estimate on D shows to what extent, on average, the share of redshirted girls differs from that of non-redshirted girls'.
Source of data: Survey data – Hungarian Life Course Survey.

Table 3: ***Academic Redshirting (cutoff: January 1) and School Start at 7 Due to Enrollment Cutoff Date (June 1)*, Average Characteristics of Compliers, Administrative Data – Grade 6 and Survey Data**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	cutoff x_d : *January 1*			*cutoff* x_d : *June 1*			
	for Academic Redshirting			**School Enrollment Cutoff**			
Panel A: Admin. Test Score Data	**share among**		**p-value** $\triangle$	**share among**		**p-value** $\triangle$	
	compliers	**sample**	**(1)-(2)**	**compliers**	**sample**	**(4)-(5)**	**(1)-(4)**
gender: boy	0.5611	0.4923	0.000***	0.4101	0.4941	0.000***	0.000***
parental education: primary	0.0982	0.1040	0.130	0.0857	0.1042	0.000***	0.002***
parental education: vocational	0.2648	0.2707	0.205	0.2832	0.2714	0.006***	0.000***
parental education: secondary	0.3173	0.3124	0.192	0.3308	0.3115	0.000***	0.009***
parental education: tertiary	0.2989	0.2945	0.477	0.2836	0.2943	0.052*	0.005***
family disadvantaged	0.2873	0.2952	0.178	0.2824	0.2943	0.030**	0.391
subsidized meals at school	0.2683	0.2715	0.343	0.2546	0.2733	0.000***	0.003***
free meals at school	0.2643	0.2561	0.140	0.2330	0.2514	0.000***	0.000***
free textbooks at school	0.4835	0.4921	0.115	0.4698	0.4919	0.000***	0.018**
only 0-50 books at home	0.1655	0.1668	0.753	0.1456	0.1657	0.000***	0.000***
own books at home	0.9461	0.9468	0.747	0.9587	0.9471	0.000***	0.000***
own desk at home	0.9186	0.9162	0.326	0.9260	0.9168	0.000***	0.022**
family considered poor	0.1616	0.1573	0.156	0.1432	0.1571	0.000***	0.000***
neighbourhood considered poor	0.1173	0.1212	0.254	0.1158	0.1204	0.178	0.684
number of siblings in household	1.3393	1.3675	0.017**	1.3045	1.3711	0.000***	0.010***
lives with stepmother	0.0194	0.0174	0.083*	0.0125	0.0173	0.000***	0.000***
lives in Budapest (capital)	0.1404	0.1219	0.033**	0.0989	0.1233	0.000**	0.000***
lives in North-Eastern of Hungary	0.1225	0.1297	0.464	0.1332	0.1302	0.750	0.105
Panel B: Survey Data	**share among**		**p-value** $\triangle$	**share among**		**p-value** $\triangle$	
	compliers	**sample**	**(1)-(2)**	**compliers**	**sample**	**(4)-(5)**	**(1)-(4)**
gender: boy	0.5547	0.5046	0.0043***	0.4035	0.5101	0.002***	0.002***
age in months	181.32	175.67	0.000***	162.77	173.70	0.000***	0.000***
birth happened prematurely	0.0882	0.0803	0.4987	0.0668	0.0836	0.1905	0.3205
birth happened w/ complications	0.1430	0.1437	0.9562	0.1371	0.1424	0.744	0.8178
born with low birth weight	0.0979	0.0915	0.6051	0.0390	0.0852	0.0003***	0.0084*
was breastfed for <3 months	0.2932	0.2920	0.9409	0.2597	0.2820	0.2748	0.3021
separated from mother (<1y old)	0.0317	0.0231	0.1731	0.0024	0.0222	0.0026***	0.0135**
separated from father (<1y old)	0.1536	0.1399	0.3024	0.1292	0.1417	0.4273	0.3435
said first words >1y old	0.3612	0.3128	0.0033***	0.2531	0.3096	0.0064	0.0011***
age (months): said first sentences	23.07	21.77	0.000***	20.45	21.74	0.0003***	0.000***
chronic face illness (<3y old)	0.0648	0.0476	0.029**	0.0523	0.0497	0.8107	0.4603
chronic ear illness (<3y old)	0.1436	0.1102	0.0046***	0.1053	0.1147	0.5155	0.092*
chronic asthma (<3y old)	0.0463	0.0513	0.4937	0.0485	0.0476	0.9242	0.8828
$\ominus$ *income shock to family (<3y old)*	0.0961	0.0798	0.1026	0.0747	0.0822	0.5551	0.289
problems w/ nerv.sys. (4-5y old)	0.0123	0.0068	0.0637*	0.0011	0.0063	0.1355	0.0607*
problems w/ speaking (4-5y old)	0.2178	0.1889	0.0562*	0.1349	0.1786	0.0125**	0.0042***
problems w/ cognition (4-5y old)	0.0219	0.0151	0.1604	0.0088	0.0156	0.1827	0.127
problems w/ coordination (4-5y old)	0.0126	0.0083	0.2506	0.0012	0.0079	0.1049	0.0968*
hyperactivity (4-5y old)	0.0222	0.0272	0.274	0.0132	0.0201	0.2388	0.2736
attention disorder (4-5y old)	0.0482	0.0182	0.0014***	0.0036	0.0280	0.0001***	0.0001***

Notes: [1] The relevant sample includes children who were born in the months October to March or March to August.
[2] The corresponding equation is (3). Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 6) 2008-2017, and Hungarian Life Course Survey.

Table 4: ***Academic Redshirting (cutoff: January 1) and School Start at 7 Due to Enrollment Cutoff Date (June 1)*, Average Characteristics of Compliers, By Gender, Survey Data**

	(1)	(2)	(3)	(4)	(5)
	boys		***girls***		
Panel A: cutoff x_d : January 1	**share among**		**share among**		**p-value △**
for Academic Redshirting	**compliers**	**sample**	**compliers**	**sample**	**(1)-(3)**
age in months	181.56	176.35	181.10	174.98	0.2443
birth happened prematurely	0.0777	0.0794	0.1029	0.0809	0.3412
birth happened w/ complications	0.1420	0.1511	0.1456	0.1362	0.8986
born with low birth weight	0.0858	0.0802	0.1127	0.1029	0.3162
was breastfed for <3 months	0.2840	0.3001	0.3066	0.2840	0.5219
separated from mother (<1y old)	0.0224	0.0242	0.0434	0.0221	0.1818
separated from father (<1y old)	0.1494	0.1409	0.1606	0.1384	0.7014
said first words >1y old	0.3882	0.3460	0.3280	0.2788	0.0822*
age (months): said first sentences	23.87	22.65	22.14	20.87	0.0117**
chronic face illness (<3y old)	0.0768	0.0567	0.0500	0.0385	0.1055
chronic ear illness (<3y old)	0.1482	0.1187	0.1390	0.1017	0.7201
chronic asthma (<3y old)	0.0427	0.0652	0.0509	0.0373	0.6101
⊖ *income shock to family (<3y old)*	0.0969	0.0808	0.0936	0.0784	0.8861
problems w/ nerv.sys. (4-5y old)	0.0085	0.0076	0.0172	0.0061	0.2281
problems w/ speaking (4-5y old)	0.2248	0.2209	0.2112	0.1563	0.6848
problems w/ cognition (4-5y old)	0.0209	0.0205	0.0233	0.0097	0.805
problems w/ coordination (4-5y old)	0.0078	0.0091	0.0188	0.0076	0.1735
hyperactivity (4-5y old)	0.0348	0.0286	0.0106	0.0076	0.0252**
attention disorder (4-5y old)	0.0656	0.0413	0.0241	0.0128	0.0008***
	boys		***girls***		
Panel B: cutoff x_d : June 1	**share among**		**share among**		**p-value △**
School Enrollment Cutoff	**compliers**	**sample**	**compliers**	**sample**	**(1)-(3)**
age in months	160.72	174.32	165.07	173.08	0.000***
birth happened prematurely	0.0607	0.0792	0.0747	0.0884	0.273
birth happened w/ complications	0.1394	0.1463	0.1352	0.1381	0.973
born with low birth weight	0.0382	0.0769	0.0430	0.0938	0.1523
was breastfed for <3 months	0.2547	0.2831	0.2667	0.2812	0.5452
separated from mother (<1y old)	0.0065	0.0240	0.0004	0.0203	0.5834
separated from father (<1y old)	0.1495	0.1496	0.1115	0.1335	0.3952
said first words >1y old	0.2580	0.3320	0.2492	0.2863	0.7939
age (months): said first sentences	20.96	22.57	19.97	20.88	0.4747
chronic face illness (<3y old)	0.0456	0.0584	0.0575	0.0407	0.6545
chronic ear illness (<3y old)	0.1380	0.1241	0.0765	0.1052	0.0419**
chronic asthma (<3y old)	0.0717	0.0594	0.0275	0.0354	0.0327**
⊖ *income shock to family (<3y old)*	0.0833	0.0823	0.0662	0.0812	0.7035
problems w/ nerv.sys. (4-5y old)	0.0020	0.0062	0.0009	0.0064	0.6185
problems w/ speaking (4-5y old)	0.1727	0.2116	0.1016	0.1449	0.0645*
problems w/ cognition (4-5y old)	0.0148	0.0202	0.0035	0.0108	0.4525
problems w/ coordination (4-5y old)	0.0055	0.0087	-0.0022	0.0070	0.6901
hyperactivity (4-5y old)	0.0203	0.0311	0.0058	0.0087	0.2785
attention disorder (4-5y old)	0.0032	0.0430	0.0032	0.0125	0.6849

Notes: [1] The relevant sample includes children who were born in the months October to March or March to August. [2] The corresponding equation is (3). Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Hungarian Life Course Survey.

General notes for Tables 5, 6, 7, and 8:
[1] The table shows the estimated coefficients and standard errors of the 2^{nd}-stage equation of Y on D and control variables. D is a binary variable denoting age at primary school entry (1: child entered primary school at age 7, 0: child entered primary school at age 6). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (2).
[2] The sample includes children born in the three-months window around the cutoff dates of January 1 and June 1, respectively.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
[4] "Control complier mean of Y" pertains to the estimated mean of Y_0 (the potential outcome of Y without the treatment), for children in the various *complier* subgroups, who were born in the three-months window around the cutoff dates of January 1 and June 1, respectively, estimated using the equation (A2) in Abdulkadiroglu et al. (2018).

Table 5: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on 10^{th}-Grade Standardized Mathematics and Reading Test Scores, and Grade Repetition by Grade 10, Administrative Data – Grade 10, For All and By Gender**

	School Start at 7 Due to Academic Redshirting *cutoff x_d : January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff x_d : June 1*		
outcome Y in grade 10:	**math**	**reading**	$1\{repeat\}$	**math**	**reading**	$1\{repeat\}$
Panel A: all students						
$D = 1$ {**start school at 7**}	0.1220**	0.0892	-0.0324	0.1393***	0.1809***	-0.0577***
	[0.056]	[0.056]	[0.020]	[0.030]	[0.029]	[0.011]
R^2	0.2433	0.2828	0.0644	0.2489	0.2807	0.0645
N	334,552	334,633	334,767	343,060	343,159	343,272
control complier mean of Y			*0.0942*			*0.1014*
Panel B: by gender						
	***male** students*					
$D = 1$ {**start school at 7**}	0.1960***	0.2469***	-0.0941***	0.1596***	0.2045***	-0.0743***
	[0.0733]	[0.0727]	[0.0265]	[0.0570]	[0.0552]	[0.0200]
R^2	0.2205	0.2321	0.0489	0.2359	0.2515	0.0690
N	161,464	161,479	161,565	166,681	166,709	166,778
control complier mean of Y			*0.1119*			*0.1199*
	***female** students*					
$D = 1$ {**start school at 7**}	0.0301	-0.1175	0.0468	0.1260***	0.1668***	-0.0494***
	[0.0883]	[0.0866]	[0.0289]	[0.0332]	[0.0319]	[0.0112]
R^2	0.2526	0.3010	0.0668	0.2465	0.2840	0.0672
N	173,088	173,154	173,202	176,379	176,450	176,494
control complier mean of Y			*0.0872*			*0.0581*

Note: In columns (1) and (4), the outcome is standardized mathematics test score in Grade 10; in columns (2) and (5), it is standardized reading test score in Grade 10; in columns (3) and (6), it is an indicator for having repeated a grade, by Grade 10.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

Table 6: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on Secondary School Track ad Aspirations, Administrative Data – Grade 10**

outcome: $Y = 1\{$**being in track** $j\}$ $(j = low, middle, high)$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
tracks	*low*	*middle*	*high*	*low*	*middle*	*high*
Panel A: all students						
$D = 1\{$**start school at 7**$\}$	-0.0732***	0.0168	0.0564*	-0.0576***	-0.0202	0.0778***
	[0.023]	[0.032]	[0.029]	[0.012]	[0.018]	[0.017]
R^2	0.2285	0.0601	0.2337	0.2089	0.0506	0.1880
N	334,767	334,767	334,767	343,272	343,272	343,272
control complier mean of Y	*0.2043*	*0.4084*	*0.3871*	*0.1628*	*0.4116*	*0.4255*
Panel B: by gender						
	male *students*					
$D = 1\{$**start school at 7**$\}$	-0.1193***	0.0126	0.1067***	-0.1021***	0.0223	0.0797***
	[0.0306]	[0.0403]	[0.0365]	[0.0214]	[0.0303]	[0.0273]
R^2	0.2168	0.0558	0.2077	0.2336	0.0584	0.2202
N	161,565	161,565	161,565	166,778	166,778	166,778
control complier mean of Y	*0.2320*	*0.4289*	*0.3390*	*0.2170*	*0.4471*	*0.3358*
	female *students*					
$D = 1\{$**start school at 7**$\}$	-0.0111	0.0207	-0.0096	-0.0363***	-0.0508***	0.0871***
	[0.0344]	[0.0516]	[0.0480]	[0.0125]	[0.0196]	[0.0185]
R^2	0.2256	0.0711	0.2283	0.2271	0.0715	0.2247865
N	173,202	173,202	173,202	176,494	176,494	176494
control complier mean of Y	*0.1629*	*0.3745*	*0.4625*	*0.1229*	*0.3865*	*0.4904*
outcome: $Y = 1\{$**aspiration** $j\}$ $(j =< HS, HS, HS+)$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
aspiration	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$
Panel C: all students						
$D = 1\{$**start school at 7**$\}$	-0.0258	-0.036	0.0618**	-0.0380***	-0.0442***	0.0822***
	[0.019]	[0.031]	[0.029]	[0.009]	[0.016]	[0.015]
R^2	0.1835	0.1134	0.2812	0.1830	0.1159	0.2815
N	329,424	329,424	329,424	337,665	337,665	337,665
control complier mean of Y	*0.1135*	*0.3556*	*0.5277*	*0.0842*	*0.3477*	*0.5699*
Panel D: by gender						
	male *students*					
$D = 1\{$**start school at 7**$\}$	-0.0524**	-0.0819**	0.1343***	-0.0623***	-0.038	0.1003***
	[0.0261]	[0.0396]	[0.0362]	[0.0197]	[0.0305]	[0.0277]
R^2	0.1914	0.0840	0.2544	0.1944	0.0891	0.2635
N	158,710	158,710	158,710	163,816	163,816	163,816
control complier mean of Y	*0.1359*	*0.3765*	*0.4875*	*0.2148*	*0.4484*	*0.3366*
	female *students*					
$D = 1\{$**start school at 7**$\}$	0.0101	0.023	-0.0331	-0.0237***	-0.0462***	0.0699***
	[0.0255]	[0.0479]	[0.0464]	[0.0085]	[0.0176]	[0.0173]
R^2	0.1593	0.1507	0.2858	0.1537	0.1536	0.2825
N	170,714	170,714	170,714	173,849	173,849	173,849
control complier mean of Y	*0.0803*	*0.3273*	*0.5922*	*0.1216*	*0.3870*	*0.4913*

Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

Table 7: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on Obtaining a Vocational Degree or a High School Degree, For All and By Gender**

	School Start at 7 Due to Academic Redshirting *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
outcome Y**:**	$1\{dropout\}$	$1\{VOCdegr\}$	$1\{HSdegr\}$	$1\{dropout\}$	$1\{VOCdegr\}$	$1\{HSdegr\}$
Panel A: all students						
$D = 1$ {**start school at 7**}	0.0023	-0.1704***	0.1600**	0.0057	-0.0581**	0.0513*
	[0.0569]	[0.0627]	[0.0740]	[0.0246]	[0.0249]	[0.0282]
R^2	0.1752	0.0701	0.2653	0.1857	0.0993	0.2951
N	47,536	47,536	47,536	57,395	57,395	57,395
control complier mean of Y	*0.0999*	*0.1092*	*0.7922*	*0.0936*	*0.1072*	*0.8010*
Panel B: by gender						
	male *students*					
$D = 1$ {**start school at 7**}	0.0089	-0.2132**	0.1919**	0.0363	-0.0684	0.0217
	[0.0727]	[0.0859]	[0.0973]	[0.0466]	[0.0517]	[0.0581]
R^2	0.1572	0.0581	0.2476	0.1564	0.1000	0.2801
N	22,810	22,810	22,810	27,670	27,670	27,670
control complier mean of Y	*0.1074*	*0.1476*	*0.7464*	*0.1054*	*0.1597*	*0.7380*
	female *students*					
$D = 1$ {**start school at 7**}	-0.0048	-0.1248	0.1272	-0.0089	-0.0524**	0.0656**
	[0.0849]	[0.0826]	[0.1050]	[0.0253]	[0.0231]	[0.0293]
R^2	0.1954	0.0634	0.2750	0.2173	0.0748	0.3042
N	24,726	24,726	24,726	29,725	29,725	29,725
control complier mean of Y	*0.0934*	*0.0765*	*0.8312*	*0.0855*	*0.0709*	*0.8447*

Note: In columns (1) and (4), the outcome ($1\{dropout\}$) is an indicator for dropping out from secondary school; in columns (2) and (5), the outcome is an indicator for obtaining a vocational degree in secondary school ($1\{VOCdegr\}$); in columns (3) and (6), the outcome is an indicator for obtaining a high school degree in secondary school ($1\{HSdegr\}$).
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2014, Linked to Administrative Labor Data.

Table 8: ***Effect of Higher SSA Due to Academic Redshirting*** **- LATE Estimates on Anxiety, Confidence and Being Bullied Measures at Grade 8, By Gender, Survey Data and Administrative Health Data**

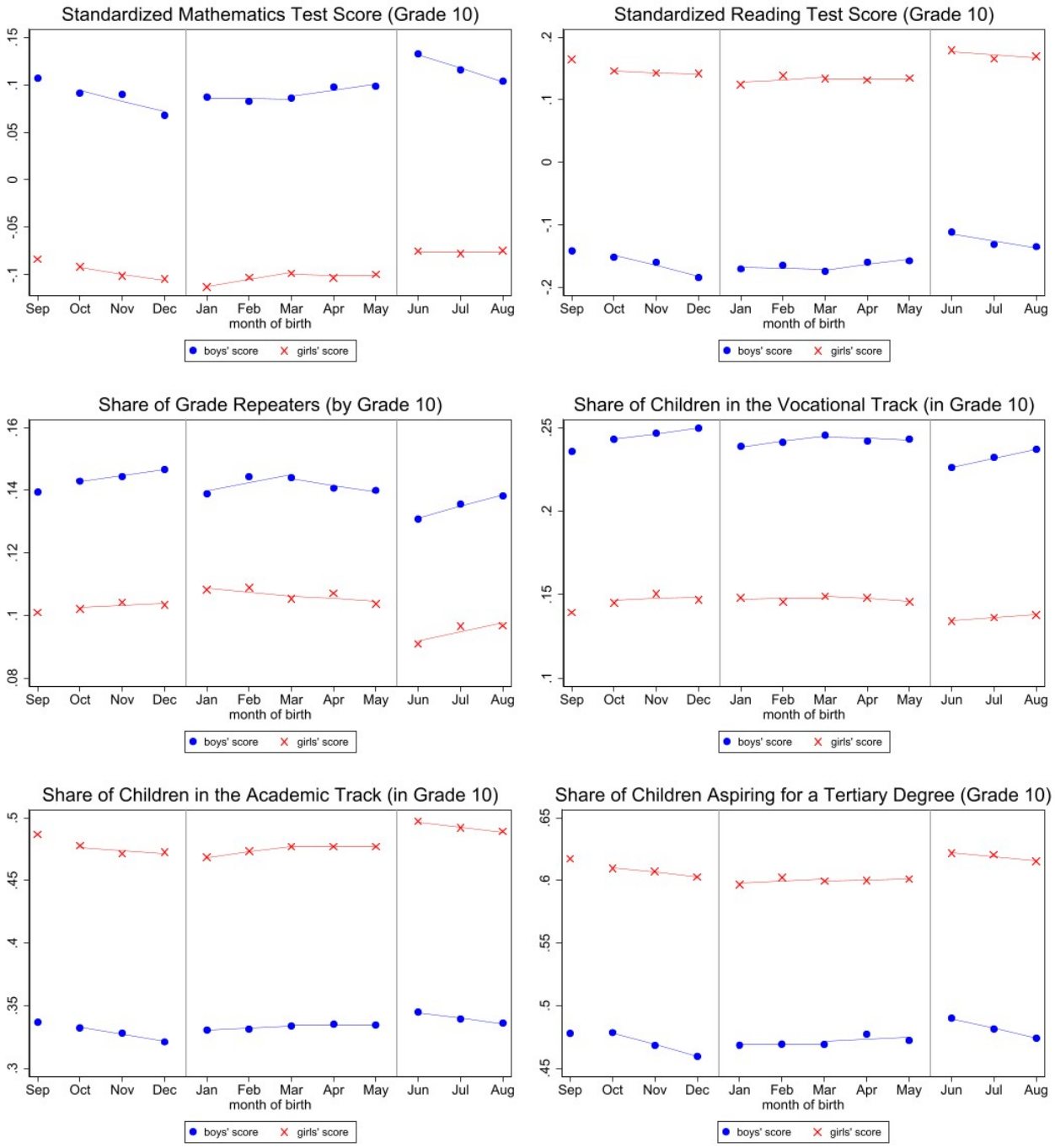
Panel A: Outcomes of Anxiety and Psycholeptics Drug Expenditures; *cutoff x_d : January 1*						
	boys			**girls**		
outcome Y:	***std. anxiety score***	***often feeling anxious***	***log(exp.) N05***	***std. anxiety score***	***often feeling anxious***	***log(exp.) N05***
$D = 1$ {**start school at 7**}	-1.1614**	-0.4555**	-0.0398**	-0.1132	-0.2905	0.0138
	[0.5101]	[0.2204]	[0.0181]	[0.6338]	[0.3212]	[0.0193]
cons	0.4360*	0.5149***	-0.01	0.5491**	0.6629***	0.0151
	[0.2508]	[0.1032]	[0.0131]	[0.2419]	[0.1210]	[0.0119]
N	38,516	38,516	62,027	37,844	37,844	64,264
1^{st}*-stage F-statistic:*	*39.44*	*39.44*	*435.54*	*27.5*	*27.5*	*406.93*
data	*survey*	*survey*	*admin health*	*survey*	*survey*	*admin health*

Panel B: Outcomes of Confidence and Being Bullied in Class; *cutoff x_d : January 1*						
	boys			**girls**		
outcome Y:	***lack of confidence***	***bullied for appearance***	***bullied for studies***	***lack of confidence***	***bullied for appearance***	***bullied for studies***
$D = 1$ {**start school at 7**}	-0.5021**	-0.3993*	-0.3272*	-0.0573	-0.166	0.0991
	[0.2476]	[0.2278]	[0.1860]	[0.3439]	[0.3348]	[0.2204]
cons	0.7604***	0.3445***	0.3732***	0.6931***	0.4465***	0.0844
	[0.1123]	[0.1041]	[0.0888]	[0.1244]	[0.1196]	[0.0880]
N	38,480	38,489	38,496	37,850	37,857	37,827
1^{st}*-stage F-statistic:*	*39.2*	*39.6*	*39.5*	*27.4*	*27.47*	*27.37*
data	*survey*	*survey*	*survey*	*survey*	*survey*	*survey*

Note: outcome Y for each column (regression) can be seen in bold in the column titles and they are:
— *standardized anxiety score* is the (standardized) mean score from students' responses to how often they experience indicators of Generalized Anxiety Disorder (GAD) in Grade 8 (score 5: 'daily'; 4: 'several times per week'; 3: 'weekly'; 2: 'monthly'; 1: 'never or rarely'): (i) having stomachache, (ii) being in bad mood and fatigued, (iii) being irritable, (iv) being fearful, (v) being nervous, (vi) having problems with falling asleep, (vii) waking up during the night, (viii) being dizzy, (ix) being exhausted, (x) feeling nausea and vomiting;
— *often feeling anxious* is an indicator of whether the student experienced any of the GAD symptoms daily, or several times per week, in Grade 8;
— *lack of confidence* by a binary variable indicating whether the student agrees with any of the following statements, in Grade 8: (i) "*I am inclined to consider myself as an unsuccessful person without any talent.*"; (ii) "*I feel I cannot be proud of myself.*"; (iii) "*I wish I had more respect for myself.*"; (iv) "*I feel I am not good at anything.*"
— *bullied* by a binary variable for being bullied in class for appearance or poor student achievement, in Grade 8;
— *N05d expenditures* which capture prescriptions for psycholeptics, in Grade 8 (ATC code N05; antipsychotics and sedatives form the major groups in this category, primarily for treating anxiety).
Source of data: Hungarian Life Course Survey and 50 percent of the administrative test score data – Hungarian National Assessment of Basic Competences (grades 6 and 8) 2009-2017, linked with drug expenditures data.

Figures

Figure 2: **Average Student Outcomes (Test Scores, Grade Repetition, Secondary School Track, Aspirations), by Month of Birth and Gender, Administrative Data – Grade 10**



Notes: The figures show the raw reduced-form relationship between month of birth and various student outcomes, by gender, together with best fit lines in the 3-months window around the January 1 and June 1 cutoffs. January 1 is the cutoff relevant for academic redshirting, and June 1 is the school enrollment cutoff date.
Source of data: Hungarian National Assessment of Basic Competences (grade 10), 2008-2017.

Appendix

A Additional Institutional Details

Hungary has a universal childcare system, with a minimal (less than 4 percent) fraction of private institutions. Public childcare institutions provide an 8-hour long service per day free of charge, except for meal fees. According to the Act No. LXXIX of 1993 on Public Education (24(1)), childcare institutions educate children from age 3 until the start of compulsory school. According to the enactment of the Ministry of Education, 137/1996. (VIII. 28.), childcare teachers have to satisfy the physical and mental needs of children: to assure a healthy lifestyle, to provide emotional security, socialization, and to help mental skill enhancement.

According to Vago (2005, pp.744), a flexible system of primary school start in Hungary came into effect in 1986,[36] with the intention to provide opportunity for the less developed and immature children to spend one additional year in childcare, while for the more developed and mature children to start school earlier than required. There were strong social and institutional reasons, other than child abilities around school entry, for redshirting. In the middle of the 1980s, Hungary started open up and for many people studying and working abroad became possible; the increasing entrepreneurial opportunities and high unemployment rates among the unskilled and young around the transition from socialism incentivized primarily the middle-class parents to provide better quality, rather than fast education for their children. Moreover, elite primary schools required students to pass increasingly difficult entrance exams that were (more) easily taken by mature children.

The physical, psychological and social requirements for school-readiness are described in the enactment of the Ministry of Education, 137/1996. (VIII. 28.), VI./2. Children meeting the physical requirements can be characterized by well-proportioned body, more and more harmonious movements, proper audition, developing coordinational skills and the ability of controlling motion, behavior and physical needs. Children meeting the psychological requirements are open to enter primary school with gradually enhancing perceptional and remembering skills, being able to pay attention to others and knowing the primary behavioral requirements, knowing the own body and primary information about herself and her environment. Children meeting the social requirements are ready to accept the school life and primary school teacher, gradually become able to cooperate, to initiate contact and to conform to the rules. The aforementioned ones are generally needed for heightened ability to learn in school. For instance, proper audition and vision is essential to learn how to read and write properly. Knowing the own body is essential, since individuals generally determine different object of the world relative to themselves. There is also some "knowledge stock" children need to have at age primary school start, or it considerably helps them later. The developmental experts highlighted the following ones in general: own name, age, parents' name and occupation, siblings' name and age, different parts of the body and their role, assessing distance ("near" and "close"), assessing the order of events, different colors, different parts of the day, different seasons, the relationship between weather and clothing, organs and senses, basic counting abilities ("more" and "less", "shorter" and longer", "whole" and "part"), *etc.*

[36] Vago, I. (2005). Felfele Terjeszkedo Ovodaztatas - Stagnalo Hozzaferes. *Educatio*, 4:742-760.

B First-Stage Results on Higher School-Starting Age

Table B1: ***Effect of Quarter of Birth* - First-Stage Results on Higher SSA (1: Starting School at the Age of 7 (vs. 6)), Due to Academic Redshirting and Due to the School Enrollment Cutoff Date, Administrative Data – Grades 6/8/10, For All and By Gender**

outcome: $D = 1\{$***start school at 7***$\}$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
grades:	**6**	**8**	**10**	**6**	**8**	**10**
Panel A: all students						
$Z = 1\{$**birth date** $X \geq x_d\}$	0.1241***	0.1243***	0.1119***	0.1703***	0.1923***	0.2141***
	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.004]
R^2	0.1761	0.1679	0.1412	0.3117	0.3243	0.3389
N	357,280	341,964	334,767	375,276	356,691	343,272
baseline mean of Y	*0.1157*	*0.0986*	*0.0929*	*0.5703*	*0.5417*	*0.5010*
Panel B: by gender						
	male *students*					
$Z = 1\{$**birth date** $X \geq x_d\}$	0.1388***	0.1413***	0.1307***	0.1233***	0.1451***	0.1665***
	[0.0044]	[0.0045]	[0.0043]	[0.0045]	[0.0047]	[0.0052]
R^2	0.1809	0.1750	0.1558	0.2720	0.2811	0.2949
N	175,557	165,514	161,565	185,021	173,806	166,778
baseline mean of Y	*0.1464*	*0.1256*	*0.1133*	*0.6410*	*0.6138*	*0.5715*
	female *students*					
$Z = 1\{$**birth date** $X \geq x_d\}$	0.1096***	0.1079***	0.0941***	0.2164***	0.2385***	0.2603***
	[0.0038]	[0.0036]	[0.0036]	[0.0051]	[0.0052]	[0.0056]
R^2	0.1533	0.1449	0.1135	0.3416	0.3562	0.3720
N	181,723	176,450	173,202	190,255	182,885	176,494
baseline mean of Y	*0.0862*	*0.0733*	*0.0739*	*0.5009*	*0.4730*	*0.4346*

Notes: [1] The table shows the estimated coefficients and standard errors of a Linear Probability Model (LPM) of D on Z and control variables. D is a binary variable denoting age at primary school entry (1: child entered primary school at age 7, 0: child entered primary school at age 6). Z is a binary variable denoting quarter of birth (1: child was born on/after January 1 or June 1, 0: child was born before January 1 or June 1). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (1).
[2] The sample includes children born in the three-months window around the cutoff dates of January 1 and June 1, respectively.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
[4] "Baseline mean of Y" pertains to mean of D_i, for children in the various subgroups, who were born in the three-months window around the cutoff dates of January 1 and June 1, respectively, for whom also $Z_i = 0$.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grades 6,8,10) 2008-2017.

C Sample Selection, Descriptives, and Evidence on Remediation

Table C1: **Details of Sample Selection, Administrative Data – Grade 10**

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
All [N]	107,654	100,620	96,898	94,047	93,167	93,907	92,544	94,127	93,214	92,198
info on: yrs in childcare	80.83%	80.48%	83.12%	82.94%	84.54%	84.38%	84.15%	82.24%	84.72%	85.94%
info on: birthdate	79.82%	80.43%	82.96%	82.80%	82.72%	84.04%	84.06%	82.43%	84.54%	82.74%
has valid test score	79.69%	80.38%	82.73%	82.62%	82.11%	83.77%	83.71%	82.19%	83.12%	80.28%
info on: parental educ.	76.54%	77.84%	80.03%	80.97%	81.94%	81.11%	81.29%	81.23%	82.12%	79.94%

Notes: [1] Each cell in the $2^{nd} - 5^{th}$ row shows the fraction (%) of sample remaining after keeping children who have information on how many years they spent in childcare, who have valid information on their year and month of birth, and who have at least one valid test score, respectively; the 4^{th} row corresponds to the *analysis sample*.
[2] The last row shows the shows sample after keeping children who have valid information about their parent's education.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

Table C2: **Share of Boys and Children with Different Parental Education, by Month of Birth and School-Starting Age, Administrative Data – Grade 10**

Month of Birth:	**October-December**		**January-March**	
cutoff x_d : January 1	**Starting at 6**	**Starting at 7**	**Starting at 6**	**Starting at 7**
fraction of boys	47.5%	61.7%	44.1%	58.2%
(mean)	*49.4%*		*49.2%*	
fraction of parental educ: at most 8 years in school	9.3%	24.2%	8.8%	15.3%
(mean)	*11.584%*		*11.310%*	
fraction of parental educ: at most vocational training	28.7%	35.1%	28.1%	31.4%
(mean)	*29.437%*		*29.4%*	
fraction of parental educ: at most high school degree	32.4%	23.467%	32.8%	28.8%
(mean)	*31.043%*		*31.2%*	
Month of Birth:	**March-May**		**June-August**	
cutoff x_d : June 1	**Starting at 6**	**Starting at 7**	**Starting at 6**	**Starting at 7**
fraction of boys	41.2%	55.7%	35.4%	49.7%
(mean)	*49.6%*		*49.5%*	
fraction of parental educ: at most 8 years in school	9.1%	12.3%	9.8%	10.9%
(mean)	*11.4%*		*11.6%*	
fraction of parental educ: at most vocational training	28.9%	28.9%	23.8%	29.7%
(mean)	*28.9%*		*29.6%*	
fraction of parental educ: at most high school degree	32.4%	30.3%	27.7%	31.5%
(mean)	*30.9%*		*30.9%*	

Notes for Interpretation: [1] The first row shows, e.g., that 47.5 (58.2) percent of children who were born October–December (January–March), and started school at the age of 6, are boys.
[2] The second row shows that the share of boys born in the three-month window around the January 1 cutoff, is around 49 percent on both sides of the cutoff: it is 49.4 percent for children born in October–December and 49.2 percent for those in January–March.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

In Table C4 below, **parental investments**: '**1**' is 1 if the parents help the child with their homework at home, '**2**' is 1 if the parents discuss with the child at home what happened in school, '**3**' is 1 if the parents discuss with the child at home what the child has read; **tutoring: any** is 1 if the child is attending any extra tutoring classes, **tutoring: math** is 1 if the child attends tutoring classes in math, and **tutoring: teacher** is 1 if the child says that the teacher devotes extra attention to them in class and'or outside class.

Table C3: ***Exogeneity of Quarter of Birth Around January 1*** **- "First-Stage" Results on Various Background and Early Childhood Characteristics, Administrative Data – Grade 6 and Survey Data**

	(1)	(2)	(3)	(4)	(5)	(6)
Admin. Test Score Data			***parental education***			
	boy	***primary***	***vocational***	***secondary***	***tertiary***	***disadv.***
$Z = 1\{X \geq x_d\}$	0.0025	0.0011	-0.0037	-0.001	0.0032	-0.0022
	[0.004]	[0.002]	[0.003]	[0.003]	[0.003]	[0.002]
R^2	0.0101	0.4030	0.1360	0.0823	0.3077	0.6516
N	369,016	369,016	369,016	369,016	369,016	369,016
	meals at school		***books at school***	***books at home***	***own desk***	***family***
	free	***supported***	***free***	***0-50***	***at home***	***poor***
$Z = 1\{X \geq x_d\}$	-0.0007	0.0013	-0.0035	0.0025	-0.0012	0.0021
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
R^2	0.3499	0.6762	0.5657	0.2198	0.1177	0.1890608
N	369,016	369,016	369,016	369,016	369,016	369016
	household	***number of***	***lives with***	***neighbourhood***	***(capital)***	***North-Eastern***
	size	***siblings***	***stepmother***	***poor***	***Budapest***	***Hungary***
$Z = 1\{X \geq x_d\}$	0.0047	0.0095	-0.0003	-0.0009	0.0016	-0.002
	[0.006]	[0.007]	[0.001]	[0.002]	[0.002]	[0.002]
R^2	0.6349	0.2830	0.0174	0.1748	0.0479	0.0238
N	368,065	369,016	369,016	369,016	368,065	369,016
Survey Data	***premature***	***birth***	***born w/ low***	***breastfed***	***separated from***	
0-1y old	***birth***	***complications***	***birth weight***	***<3 months***	***mother***	***father***
$Z = 1\{X \geq x_d\}$	-0.0076	-0.001	-0.0123	-0.0423	0.0073	-0.0098
	[0.016]	[0.021]	[0.017]	[0.029]	[0.007]	[0.019]
R^2	0.0088	0.0116	0.0496	0.0339	0.3542	0.0955
N	79,441	79,441	79,422	77,148	78,369	78,383
1-3y old	1^{st} ***words***	1^{st} ***sentences***	***chronic illness in <3y***			***income***
	> 1y old	***age in months***	***face***	***ear***	***asthma***	***shock***
$Z = 1\{X \geq x_d\}$	-0.0053	-0.2461	-0.0137	0.0028	0.0111	0.0037
	[0.027]	[0.410]	[0.013]	[0.018]	[0.012]	[0.014]
R^2	0.0217	0.0382	0.0137	0.0072	0.0137	0.0516
N	76,921	68,468	78,810	78,904	79,381	78,983
4-5y old			***problems with***		***hyper-***	***attention***
	nervsys	***speaking***	***cognition***	***coord***	***activity***	***disorder***
$Z = 1\{X \geq x_d\}$	-0.005	-0.0247	-0.0038	-0.0016	0.0016	0.0086
	[0.005]	[0.023]	[0.006]	[0.004]	[0.006]	[0.007]
R^2	0.0040	0.0229	0.0133	0.0048	0.0163	0.0232
N	79,427	79,182	79,375	79,387	79,343	79,365

Notes: [1] The table shows the estimated coefficients and standard errors of various outcome variables (indicated in bold in the column titles) on Z and control variables. Z is a binary variable denoting quarter of birth (1: child was born on/after January 1 or June 1, 0: child was born before January 1 or June 1). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (1).
[2] The sample includes children born in the three-months window around the cutoff dates of January 1.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017, and Hungarian Life Course Survey.

Table C4: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on Parental Involvement and Tutoring, Administrative (Grade 10) and Survey Data**

outcome: $Y = 1\{$**parents invest/help**$\}$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
parental investments	**1**	**2**	**3**	**1**	**2**	**3**
Panel A: all students						
$D = 1\{$**start school at 7**$\}$	-0.008	-0.0124	-0.0064	-0.0066	-0.0480***	-0.0006
	[0.017]	[0.031]	[0.018]	[0.009]	[0.017]	[0.009]
R^2	0.007	0.048	0.008	0.008	0.045	0.008
N	329,949	334,767	334,767	338,098	343,272	343,272
control complier mean of Y	*0.0723*	*0.5872*	*0.0867*	*0.0607*	*0.5970*	*0.0728*
Panel B: by gender						
	male *students*					
$D = 1\{$**start school at 7**$\}$	0.0063	0.0455	-0.0115	0.0001	-0.0807**	0.0169
	[0.0222]	[0.0402]	[0.0212]	[0.0169]	[0.0318]	[0.0173]
R^2	0.006	0.043	0.008	0.007	0.040	0.009
N	158,720	161,565	161,565	163,701	166,778	166,778
control complier mean of Y	*0.0770*	*0.5455*	*0.0905*	*0.0680*	*0.5586*	*0.0733*
	female *students*					
$D = 1\{$**start school at 7**$\}$	-0.0275	-0.0866*	-0.0006	-0.0106	-0.0288	-0.0114
	[0.0265]	[0.0492]	[0.0298]	[0.0102]	[0.0194]	[0.0106]
R^2	0.004	0.036	0.009	0.007	0.038	0.008
N	171,229	173,202	173,202	174,397	176,494	176,494
control complier mean of Y	*0.0657*	*0.6506*	*0.0805*	*0.0550*	*0.6270*	*0.0725*
outcome: $Y = 1\{$**tutoring**$\}$ $(j =< HS, HS, HS+)$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*			***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*		
tutoring	**any**	**math**	**teacher**	**any**	**math**	**teacher**
Panel C: all students						
$D = 1\{$**start school at 7**$\}$	-0.0025	0.018	0.0289	-0.015	-0.0018	-0.0284
	[0.029]	[0.027]	[0.2034]	[0.014]	[0.013]	[0.0886]
R^2	0.008	0.032	0.016	0.008	0.030	0.020
N	319,616	323,986	66,488	327,456	332,096	59,495
control complier mean of Y	*0.2346*	*0.2082*	*0.5451*	*0.1917*	*0.1781*	*0.5202*
Panel D: by gender						
	male *students*					
$D = 1\{$**start school at 7**$\}$	0.0178	0.0189	-0.1166	0.0027	-0.0043	-0.1251
	[0.0352]	[0.0303]	[0.2665]	[0.0264]	[0.0240]	[0.1434]
R^2	0.010	0.034	0.017	0.009	0.032	0.015
N	152,944	155,533	33,457	157,712	160,448	30,303
control complier mean of Y	*0.2158*	*0.1960*	*0.5782*	*0.1797*	*0.1568*	*0.5472*
	female *students*					
$D = 1\{$**start school at 7**$\}$	-0.0281	0.0174	0.1977	-0.0263	-0.0003	0.0784
	[0.0472]	[0.0472]	[0.3337]	[0.0165]	[0.0163]	[0.1165]
R^2	0.006	0.027	-0.002	0.007	0.027	0.027
N	166,672	168,453	32,897	169,744	171,648	29,094
control complier mean of Y	*0.2596*	*0.2278*	*0.5109*	*0.1984*	*0.1909*	*0.5051*

Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017, and Hungarian Life Course Survey.

D LATE Estimates on Student Achievement in Primary School

Table D1: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on 6^{th}-Grade and 8^{th}-Grade Standardized Mathematics and Reading Test Scores, Administrative Data – Grades 6 and 8, For All and By Gender**

outcome:	***School Start at 7 Due to***		***School Start at 7 Due to***	
Y: std. mathematics score	***Academic Redshirting***		***Enrollment Cutoff Date***	
	cutoff x_d : *January 1*		*cutoff* x_d : *June 1*	
in grades:	**6**	**8**	**6**	**8**
Panel A: all students				
$D = 1\{$**start school at 7**$\}$	0.2124***	0.1448***	0.2520***	0.1482***
	[0.051]	[0.054]	[0.039]	[0.033]
R^2	0.2057	0.2226	0.2203	0.2330
N	357,158	341,803	375,163	356,517
Panel B: by gender				
	male *students*			
$D = 1\{$**start school at 7**$\}$	0.3015***	0.1433**	0.1934**	0.1225*
	[0.0691]	[0.0686]	[0.0757]	[0.0659]
R^2	0.1984	0.2277	0.2303	0.2385
N	175,507	165,447	184,955	173,727
	female *students*			
$D = 1\{$**start school at 7**$\}$	0.103	0.1408*	0.2836***	0.1612***
	[0.0784]	[0.0823]	[0.0418]	[0.0366]
R^2	0.2102	0.2170	0.2098	0.2258
N	181,651	176,356	190,208	182,790
outcome:	***School Start at 7 Due to***		***School Start at 7 Due to***	
Y: std. reading score	***Academic Redshirting***		***Enrollment Cutoff Date***	
	cutoff x_d : *January 1*		*cutoff* x_d : *June 1*	
in grades:	**6**	**8**	**6**	**8**
Panel C: all students				
$D = 1\{$**start school at 7**$\}$	0.2043***	0.1302**	0.2750***	0.2397***
	[0.050]	[0.051]	[0.036]	[0.034]
R^2	0.2700	0.2853	0.2830	0.2872
N	357,230	341,881	375,223	356,619
Panel D: by gender				
	male *students*			
$D = 1\{$**start school at 7**$\}$	0.2906***	0.2167***	0.2731***	0.2691***
	[0.0659]	[0.0658]	[0.0732]	[0.0661]
R^2	0.2317	0.2426	0.2575	0.2531
N	175,527	165,474	184,991	173,769
	female *students*			
$D = 1\{$**start school at 7**$\}$	0.0937	0.0203	0.2729***	0.2207***
	[0.0770]	[0.0767]	[0.0386]	[0.0355]
R^2	0.2852	0.3025	0.2881	0.2976
N	181,703	176,407	190,232	182,850

Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grades 6/8) 2008-2017.

Table D2: ***Effect of Higher SSA Due to Academic Redshirting and Due to the Enrollment Cutoff Date* - LATE Estimates on Grade Repetition by a Given Grade, Administrative Data, Grades 6 and 8, For All and By Gender**

outcome: $Y = 1\{$***having repeated a grade***$\}$	***School Start at 7 Due to Academic Redshirting*** *cutoff* x_d : *January 1*		***School Start at 7 Due to Enrollment Cutoff Date*** *cutoff* x_d : *June 1*	
by grades	**6**	**8**	**6**	**8**
Panel A: all students				
$D = 1\{$**start school at 7**$\}$	-0.0299***	-0.0144	-0.0306***	-0.0360***
	[0.011]	[0.012]	[0.009]	[0.008]
R^2	0.0826	0.0753	0.0907	0.0754
N	357,280	341,964	375,276	356,691
control complier mean of Y	*0.04383*	*0.0482*	*0.0394*	*0.0456*
Panel B: by gender				
	male *students*			
$D = 1\{$**start school at 7**$\}$	-0.0512***	-0.0209	-0.0310*	-0.0540***
	[0.0160]	[0.0166]	[0.0182]	[0.0165]
R^2	0.0834	0.0795	0.1041	0.0838
N	175,557	165,514	185,021	173,806
control complier mean of Y	*0.0501*	*0.0584*	*0.0485*	*0.0581*
	female *students*			
$D = 1\{$**start school at 7**$\}$	-0.0044	-0.0056	-0.0309***	-0.0267***
	[0.0163]	[0.0175]	[0.0088]	[0.0082]
R^2	0.0784	0.0692	0.0860	0.0744
N	181,723	176,450	190,255	182,885
control complier mean of Y	*0.0334*	*0.0335*	*0.0318*	*0.0364*

Notes: [1] The table shows the estimated coefficients and standard errors of the 2^{nd}-stage equation of Y on D and control variables. Y is a binary variable denoting whether the child repeated at least one grade by grades 6, 8 and 10, respectively (1: child repeated grade, 0: child never repeated a grade). D is a binary variable denoting age at primary school entry (1: child entered primary school at age 7, 0: child entered primary school at age 6). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (2).
[2] The sample includes children born in the three-months window around the cutoff dates of January 1 and June 1, respectively.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
[4] "Control complier mean of Y" pertains to mean of Y_{0i} (the potential outcome of Y without the treatment), for children in the various *complier* subgroups, who were born in the three-months window around the cutoff dates of January 1 and June 1, respectively, estimated using the regression equation (A2) in Abdulkadiroglu et al. (2018).
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grades 6/8) 2008-2017.

E Placebo Tests for Alternative Cutoffs Around January 1

Table E1: ***Placebo Tests for Alternative Cutoffs Around January 1*** **- First-Stage and Reduced-Form Estimates on Standardized Test Score, Grade Repetition, Secondary School Track Choice and Educational Aspirations, For All and For Boys, Administrative Data – Grade 10**

		placebo cutoffs x_d					
		Oct1	***Nov1***	***Dec1***	***Jan1***	***Feb1***	***Mar1***
		1^{st}***-stage*** // $D = 1\{$**start school at 7**$\}$					
all	$Z = 1\{X \geq x_d\}$	0.0077***	0.0123***	0.0344***	**0.1387*** **	0.0563***	0.0111**
		[0.0015]	[0.0027]	[0.0019]	[0.0026]	[0.0028]	[0.0049]
boys	$Z = 1\{X \geq x_d\}$	0.0100***	0.0163***	0.0435***	**0.1645*** **	0.0616***	0.0127*
		[0.0023]	[0.0041]	[0.0030]	[0.0038]	[0.0039]	[0.0071]
		reduced-form // Y_i***: std. mathematics score in grade 10***					
all	$Z = 1\{X \geq x_d\}$	-0.0007	-0.0039	-0.0032	**0.0097***	0.0007	0.0006
		[0.0050]	[0.0086]	[0.0052]	[0.0051]	[0.0051]	[0.0088]
boys	$Z = 1\{X \geq x_d\}$	-0.003	-0.0025	-0.0065	**0.0200*** **	-0.0015	-0.0004
		[0.0074]	[0.0129]	[0.0075]	[0.0076]	[0.0077]	[0.0134]
		reduced-form // Y_i***: std. reading score in grade 10***					
all	$Z = 1\{X \geq x_d\}$	-0.0022	-0.0036	-0.0036	**0.0026**	0.007	-0.0162*
		[0.0051]	[0.0089]	[0.0048]	[0.0049]	[0.0051]	[0.0086]
boys	$Z = 1\{X \geq x_d\}$	0.0036	-0.0175	-0.0111	**0.0168** **	0.0081	-0.0212
		[0.0075]	[0.0133]	[0.0071]	[0.0074]	[0.0076]	[0.0130]
		reduced-form // $Y_i = 1\{$***having repeated a grade by grade 10***$\}$					
all	$Z = 1\{X \geq x_d\}$	-0.0001	0.0026	-0.0019	**-0.0031***	0.0015	-0.0022
		[0.0017]	[0.0029]	[0.0018]	[0.0018]	[0.0017]	[0.0030]
boys	$Z = 1\{X \geq x_d\}$	-0.0004	0.0036	-0.0003	**-0.0104*** **	0.0015	-0.0003
		[0.0026]	[0.0044]	[0.0027]	[0.0028]	[0.0027]	[0.0048]
		reduced-form // $Y_i = 1\{$**being in track** "*vocational*"$\}$					
all	$Z = 1\{X \geq x_d\}$	0.0021	0.0029	-0.0037*	**-0.0057*** **	0.0009	0.0032
		[0.0020]	[0.0036]	[0.0020]	[0.0019]	[0.0019]	[0.0033]
boys	$Z = 1\{X \geq x_d\}$	0.0013	0.0036	-0.0008	**-0.0119*** **	0.0012	0.005
		[0.0031]	[0.0057]	[0.0031]	[0.0030]	[0.0031]	[0.0052]
		reduced-form // $Y_i = 1\{$**being in track** "*academic*"$\}$					
all	$Z = 1\{X \geq x_d\}$	-0.0026	-0.0009	-0.0003	**0.004**	0.0012	0.0005
		[0.0026]	[0.0044]	[0.0026]	[0.0026]	[0.0027]	[0.0048]
boys	$Z = 1\{X \geq x_d\}$	-0.0002	-0.0028	-0.0048	**0.0106*** **	0.0001	0.0018
		[0.0038]	[0.0065]	[0.0037]	[0.0037]	[0.0036]	[0.0064]
		reduced-form // $Y_i = 1\{$**aspiring for** $HS+$ **(in grade 10)**$\}$					
all	$Z = 1\{X \geq x_d\}$	0.0007	-0.0053	-0.0036	**0.0018**	0.0011	-0.0046
		[0.0025]	[0.0044]	[0.0026]	[0.0026]	[0.0027]	[0.0048]
boys	$Z = 1\{X \geq x_d\}$	0.0061	-0.0154**	-0.0053	**0.0084** **	0.0021	-0.0059
		[0.0037]	[0.0065]	[0.0038]	[0.0038]	[0.0039]	[0.0067]

Notes: [1] The table shows the estimated coefficients and standard errors of the 1^{st}-stage and reduced form equations, when using a 1-month window around a particular cutoff x_d; for instance, column (1) uses only children who were born in either September or in October. Z is a binary variable denoting month of birth (1: child was born on/after a given cutoff x_d, 0: child was born before the given cutoff x_d). Control variables include: cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (1).
[2] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

F Robustness Check: LATE of Higher SSA Due to Redshirting; Window Length and Functional Form around January 1

Table F1: ***Robustness Check with a 1–2–3-Months Window and a Linear/Quadratic Trend Around January 1*** **- LATE Estimates of SSA Due to Academic Redshirting on Standardized Test Score, Grade Repetition, Secondary School Track Choice and Educational Aspirations, All Children, Administrative Data – Grades 6/8/10**

	1-M Window, No Trend			***2-M Window, Lin. Trend***			***3-M Window, Quadr. Trend***		
Y***: mathematics test score***									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	0.0868***	0.0711**	0.0702*	0.2088***	0.1782**	0.1225	0.2809**	0.2572**	0.1303
	[0.0315]	[0.0340]	[0.0367]	[0.0776]	[0.0838]	[0.0857]	[0.1097]	[0.1172]	[0.1198]
R^2	0.218	0.227	0.245	0.204	0.215	0.243	0.196	0.207	0.242
N	121,641	116,361	114,077	235,839	225,493	221,062	357,158	341,803	334,552
Y***: reading test score***									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	0.0792**	0.0540*	0.0187	0.1974***	0.1500*	0.057	0.2636**	0.2491**	0.2134*
	[0.0308]	[0.0326]	[0.0356]	[0.0756]	[0.0796]	[0.0796]	[0.1062]	[0.1164]	[0.1133]
R^2	0.279	0.291	0.284	0.269	0.281	0.285	0.263	0.271	0.268
N	121,662	116,376	114,106	235,889	225,535	221,114	357,230	341,881	334,633
$Y = 1\{$***having repeated a grade***$\}$									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	-0.0194***	-0.0094	-0.0220*	-0.0263	-0.0189	-0.0109	-0.0096	-0.0241	0.0139
	[0.0072]	[0.0081]	[0.0129]	[0.0180]	[0.0197]	[0.0302]	[0.0253]	[0.0281]	[0.0412]
R^2	0.082	0.077	0.064	0.083	0.074	0.067	0.087	0.074	0.068
N	121,682	116,402	114,155	235,926	225,591	221,208	357,280	341,964	334,767
$Y = 1\{$**being in track** $j\}$									
track	*middle*	*high*	*low*	*middle*	*high*	*low*	*middle*	*high*	*low*
D	-0.0412***	0.0123	0.0289	-0.0184	-0.0236	0.042	-0.0394	0.0026	0.0369
	[0.0140]	[0.0209]	[0.0188]	[0.0328]	[0.0466]	[0.0427]	[0.0480]	[0.0664]	[0.0637]
R^2	0.231	0.061	0.235	0.238	0.061	0.235	0.236	0.060	0.236
N	114,155	114,155	114,155	221,208	221,208	221,208	334,767	334,767	334,767
$Y = 1\{$**aspiration** $j\}$									
level	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$
D	-0.014	0.0008	0.0132	0.0043	-0.0577	0.0534	-0.0281	-0.0757	0.1038
	[0.0124]	[0.0206]	[0.0190]	[0.0287]	[0.0466]	[0.0432]	[0.0408]	[0.0703]	[0.0648]
R^2	0.182	0.115	0.284	0.187	0.111	0.281	0.183	0.109	0.274
N	112,368	112,368	112,368	217,784	217,784	217,784	329,424	329,424	329,424

Notes: [1] The table shows the estimated coefficients and standard errors of the 2^{nd}-stage equation of Y on D and control variables. D is a binary variable denoting age at primary school entry (1: child entered primary school at age 7, 0: child entered primary school at age 6). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (2).
[2] The sample includes children born in the two- vs. three-months window around the cutoff dates of January 1.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grades 6,8,10) 2008-2017.

Table F2: ***Robustness Check with a 1–2–3-Months Window and a Linear/Quadratic Trend Around January 1*** **- LATE Estimates of SSA Due to Academic Redshirting on Standardized Test Score, Grade Repetition, Secondary School Track Choice and Educational Aspirations, Boys, Administrative Data – Grades 6/8/10**

	1-M Window, No Trend			***2-M Window, Lin. Trend***			***3-M Window, Quadr. Trend***		
Y***: mathematics test score***									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	0.1601***	0.0810*	0.1216***	0.3345***	0.1647	0.2170**	0.3716**	0.2677*	0.2467
	[0.0422]	[0.0431]	[0.0464]	[0.1098]	[0.1124]	[0.1071]	[0.1447]	[0.1616]	[0.1559]
R^2	0.217	0.233	0.229	0.190	0.222	0.218	0.186	0.210	0.212
N	59,790	56,272	55,230	115,888	109,295	106,767	175,507	165,447	161,464
Y***: reading test score***									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	0.1483***	0.1322***	0.1021**	0.2894***	0.3095***	0.2242**	0.3299**	0.4369***	0.4747***
	[0.0401]	[0.0421]	[0.0452]	[0.1056]	[0.1085]	[0.1024]	[0.1391]	[0.1619]	[0.1576]
R^2	0.245	0.250	0.244	0.229	0.225	0.234	0.188	0.168	0.155
N	59,791	56,276	55,242	115,898	109,306	106,775	181,633	170,849	166,273
$Y = 1\{$***having repeated a grade***$\}$									
grade	**6**	**8**	**10**	**6**	**8**	**10**	**6**	**8**	**10**
D	-0.0326***	-0.0171	-0.0634***	-0.0412	0.0003	-0.0845**	-0.0215	0.025	-0.0882
	[0.0098]	[0.0106]	[0.0171]	[0.0259]	[0.0275]	[0.0397]	[0.0343]	[0.0398]	[0.0550]
R^2	0.087	0.081	0.055	0.086	0.082	0.052	0.092	0.080	0.051
N	59,806	56,290	55,272	115,923	109,334	106,836	175,557	165,514	161,565
$Y = 1\{$**being in track** $j\}$									
track	*middle*	*high*	*low*	*middle*	*high*	*low*	*middle*	*high*	*low*
D	-0.0723***	0.0077	0.0646***	-0.0905**	-0.0359	0.1263**	-0.1359**	0.0075	0.1284
	[0.0184]	[0.0249]	[0.0226]	[0.0445]	[0.0586]	[0.0529]	[0.0660]	[0.0864]	[0.0791]
R^2	0.225	0.058	0.212	0.223	0.056	0.202	0.211	0.056	0.203
N	55,272	55,272	55,272	106,836	106,836	106,836	161,565	161,565	161,565
$Y = 1\{$**aspiration** $j\}$									
level	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$
D	-0.028	-0.0232	0.0512**	-0.0062	-0.1075*	0.1137**	-0.0649	-0.1081	0.1729**
	[0.0171]	[0.0258]	[0.0230]	[0.0402]	[0.0600]	[0.0540]	[0.0573]	[0.0895]	[0.0816]
R^2	0.192	0.088	0.265	0.197	0.081	0.258	0.188	0.080	0.244
N	54,322	54,322	54,322	105,045	105,045	105,045	158,710	158,710	158,710

Notes: [1] The table shows the estimated coefficients and standard errors of the 2^{nd}-stage equation of Y on D and control variables. D is a binary variable denoting age at primary school entry (1: child entered primary school at age 7, 0: child entered primary school at age 6). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls. The corresponding regression equation is (2).
[2] The sample includes children born in the two- vs. three-months window around the cutoff dates of January 1.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grades 6,8,10) 2008-2017.

G Robustness Check: Standard Error Estimation

Table G1: ***Robustness Check*** **- Standard Error Estimates and t-Statistics of Coefficient Estimates on Variable $D = 1\{\text{start school at 7}\}$, for Alternative Ways of Estimating Standard Errors, Administrative Data – Grade 10, For All and for Boys**

		clustering on:						**estimates with**		
		school ID		***birth month***		***(birthm)×yr***		**'rdhonest'**		
outcome//(sub)group	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Y: std. math score	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	0.1220	0.0564	2.16	0.0029	42.65	0.0561	2.18	0.1806	0.0660	2.82
boys	0.1960	0.0733	2.67	0.0108	18.11	0.0720	2.72	0.2837	0.0846	3.35
Y: std. reading score	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	0.0892	0.0558	1.60	0.0262	3.41	0.0537	1.66	0.1440	0.0658	2.19
boys	0.2469	0.0727	3.40	0.0320	7.73	0.0706	3.50	0.3138	0.0836	3.76
$Y = 1\{repeat\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	-0.0324	0.0199	-1.63	0.0096	-3.39	0.0185	-1.75	-0.0318	0.0200	-1.59
boys	-0.0941	0.0265	-3.55	0.0034	-28.05	0.0243	-3.87	-0.0930	0.0265	-3.51
$Y = 1\{track : low\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	-0.0732	0.0233	-3.14	0.0194	-3.78	0.0217	-3.37	-0.0936	0.0257	-3.64
boys	-0.1193	0.0306	-3.9	0.0113	-10.54	0.0267	-4.47	-0.1479	0.0349	-4.24
$Y = 1\{track : high\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	0.0564	0.0290	1.94	0.0059	9.52	0.0264	2.14	0.0797	0.0328	2.42
boys	0.1067	0.0365	2.93	0.0070	15.2	0.0283	3.77	0.1373	0.0391	3.51
$Y = 1\{asp :< HS\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	-0.0258	0.0190	-1.36	0.0110	-2.36	0.0152	-1.70	-0.0381	0.0201	-1.89
boys	-0.0524	0.0261	-2.01	0.0169	-3.1	0.0162	-3.22	-0.0686	0.0285	-2.41
$Y = 1\{asp : HS+\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$
all	0.0618	0.0293	2.11	0.0085	7.3	0.0230	2.69	0.0887	0.0336	2.64
boys	0.1343	0.0362	3.71	0.0113	11.87	0.0358	3.75	0.1657	0.0420	3.95
$Z = 1\{X \geq x_d\}$	$\hat{coeff}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$	$\hat{SE}$	$\hat{t}$			
all	0.1119	0.0029	38.58	0.0045	24.86	0.0119	9.42			
boys	0.1307	0.0043	30.29	0.0057	23.03	0.0134	9.77			

Notes: [1] This table shows the standard error estimates and corresponding value of t-statistics, of coefficient estimates on variable $D = 1\{\text{start school at 7}\}$, for alternative ways of estimating standard errors (different level of clustering, as well as with the 'rdhonest' STATA command implementing Kolesar and Rothe (2018), for the case when the running variable only takes a moderate number of distinct values – in this case, the month of birth takes on just 6 values).
[2] Columns (1)–(3) show the baseline results, already presented in Tables 5 and 6.
[3] Columns (4)–(5) show the standard error estimates and t-statistics if clustering was at the level of birth-month ($N_{clusters} = 6$) and Columns (6)–(7) show them if clustering was at the level of (birth-month)×year ($N_{clusters} = 60$).
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

H Robustness Check: LATE of Higher SSA, by Accounting for Childcare-Starting Age (CSA) and Years Spent in Childcare ($yearsCC$)

The estimated effects presented in Sections 7.1 and 7.2 encompass the fact that children who start school a year older typically also start childcare a year later. Moreover, children around the January 1 and June 1 cutoffs may differ in their propensities to start childcare at various ages, affecting the length of childcare. Then, one might wonder to what extent results presented in Sections 7.1 and 7.2 are robust to accounting for the potentially endogenous childcare-starting age (CSA) and years spent in childcare ($yearsCC$). Given the identity of $SSA = CSA + yearsCC$ holds, in this Appendix, I jointly estimate the effect of SSA and $yearsCC$, where SSA is still captured by an indicator D_i that is 1 if the child started school at the age of 7 (*vs.* 6).

First, in Table H1 below, I show that the share of children who start childcare at the age of 3 or later jumps by 19 percentage points at the January 1 cutoff, which comes from the fact that 28.7 percent of December-born children start childcare at the age of 2, but only 9.3 percent of January-born children do so. (Note that there is no jump in this share at the June 1 cutoff.) At the same time, the share of children who start childcare at the age of 4 or later jumps by only 4.5 percentage points at the January 1 cutoff, and jumps by 12 percentage points at the June 1 cutoff. The average years spent in childcare is 0.092 (3.146 − 3.238) *lower* among January-born children than among December-born ones, and it is 0.158 (3.498 − 3.340) *higher* among June-born children than among May-born ones. Thus, the cutoffs do influence to some extent years spent in childcare ($yearsCC$).

To jointly estimate the effect of D and $yearsCC$, I present a variant of estimation equation (2), in which I include and instrument for $yearsCC$. Note that it is ambiguous in what order parents make decisions about their children's educational path. They might decide sequentially, first deciding about childcare starting age (CSA), and only at the relevant school-starting age (SSA) do they decide about redshirting, determining $yearsCC$. However, they might very well be forward-looking, and decide about CSA with the possibility of redshirting already in mind. If the latter is the case, then the less time spent in childcare is part of the mechanism through which redshirting impacts child outcomes, and is ideally not controlled for. Given the ambiguity regarding the timing of parental decisions, I present the regression that accounts for CSA as a variant, and not as the main specification (but, the two specifications yield very similar estimates for the impact of redshirting).

To account for the endogenous nature of $yearsCC$, I use the fraction of 3-year-old children in childcare in the child's municipality at her age of 2 as an instrument, for the following institutional reason: as described in Section 3, September-to-December-born children are more likely to enter childcare at age 2, as they are preferred by the childcare institutions if there is a queue for childcare by 2-year-olds.[37] Thus, the fraction of 3-year-old children in childcare is strongly related to whether a 2-year-old can enter.[38]

Table H2 shows the result of the first-stage regressions on the variable D_i, that is 1 if the child entered school at the age of 7, as opposed to 6, and years spent in childcare, $yearsCC_i$, for children born in a 3-months window around the January 1 and June 1 cutoffs, respectively. Regarding $yearsCC_i$, for both boys and girls, the estimated coefficients on the fraction of 3-year-old children in childcare (in the child's municipality) at her age of 2 are in line with as predicted by the institutional background. On the one hand, a higher fraction of 3-year-olds is associated with more years in childcare (due to a lower likelihood of entering at age 3+) for children born before January 1; these children turn 3 years old within 6 months of acceptance, and thus are preferred to enter at age 2 if local supply constraints permit. On the other hand, a higher fraction of 3-year-olds is associated with a significantly lower increase in years of childcare for children born after January 1; these children are not preferred to enter at age 2, if there is queuing for childcare at that age. The *joint F-statistic* is above 80 for both D_i and CC_i, and the values of both of the *Cragg-Donald Wald F-statistics* and the *Kleibergen-Papp F-statistics* are above 80 as well. Therefore, weak identification is not a concern.

Table H3 reveals that the effect of starting school a year older, due to academic redshirting using the January 1 cutoff, is remarkably robust for both boys and girls, when accounting for $yearsCC$: fixing it, redshirted boys achieve a significantly higher score in math and reading, on average, by 0.6 and 0.275 SDs, on Grade 10 testings, respectively, and they are 7.4 percentage points less likely than on-time children to repeat a grade by Grade 10;

[37] In particular, the childcare institution can admit 2-year-old children, provided that all the 3-year-old and older applicants living in the particular municipality are admitted (Public Education Act 24(1)).

[38] I construct the instrument the following way. First, for each municipality, I collect the number of 3-years-old children in childcare from KIR-Stat, and aggregate institution-level data to municipality-level. /KIR-Stat provides the most comprehensive data about the Hungarian educational system; all institutions all years are required to fill out a data form according to the enactment of the Ministry of Education 229/2006. (XI. 20.)./ Second, for each municipality, I collect the total number of 3-years-old children from the municipality-level demographic data set by cohorts of the Hungarian Statistical Authority. Third, I form the fraction of 3-years-old children in childcare and attach this measure to each child given the municipality the child lived in at her age of 2, by assuming that her family has not moved since.

they are 15 percentage points less likely to attend the lowest, vocational, high school track, and 13.5 percentage points more likely to attend the highest, academic high school track; they are 6 (13) percentage points less (more) likely to aspire to educational levels that are lower (higher) than obtaining a high school degree. For girls, we continue to see no significantly positive effects of redshirting. Interestingly, the significantly positive effects of starting school a year older due to being born after and complying with the school enrollment cutoff date of June 1 typically disappears when accounting for *yearsCC*. The *Hansen J-statistics* are typically below 1, and never higher than 2.6, thus the null hypothesis of valid (excludable) instruments cannot be rejected at the usual sig. levels. My results are in line with those of Szabo-Morvai et al. (2023).

Table H1: **Distribution, Fraction of Children Entering School at Age 7 and Average Years Spent in Childcare, by Month of Birth and Childcare-Starting Age (CSA), Administrative Data – Grade 10**

month of birth	childcare-starting age (CSA)	2	3	4	5	average
September	**frequency**	**48.29%**	**39.56%**	**8.69%**	**3.46%**	
	fraction of children entering school at age 7	*0.89%*	*7.71%*	*26.80%*	*29.52%*	6.62%
	average years spent in childcare	3.987	3.078	2.316	1.736	3.404
October	**frequency**	**43.16%**	**43.93%**	**9.21%**	**3.69%**	
	fraction of children entering school at age 7	*1.26%*	*7.84%*	*28.80%*	*31.38%*	7.57%
	average years spent in childcare	3.992	3.079	2.337	1.729	3.355
November	**frequency**	**38.03%**	**47.95%**	**9.88%**	**4.14%**	
	fraction of children entering school at age 7	*1.65%*	*9.30%*	*33.35%*	*35.48%*	9.55%
	average years spent in childcare	3.996	3.094	2.385	1.766	3.312
December	**frequency**	**28.66%**	**55.62%**	**11.14%**	**4.57%**	
	fraction of children entering school at age 7	*2.78%*	*12.23%*	*38.41%*	*39.89%*	13.38%
	average years spent in childcare	4.001	3.124	2.437	1.794	3.238
January	**frequency**	**9.27%**	**70.50%**	**14.18%**	**6.05%**	
	fraction of children entering school at age 7	*10.74%*	*23.25%*	*56.36%*	*56.43%*	28.22%
	average years spent in childcare	4.040	3.233	2.635	1.961	3.146
February	**frequency**	**7.19%**	**71.28%**	**15.32%**	**6.21%**	
	fraction of children entering school at age 7	*13.98%*	*28.79%*	*61.06%*	*60.24%*	34.10%
	average years spent in childcare	4.046	3.289	2.685	1.965	3.169
March	**frequency**	**5.50%**	**71.21%**	**16.52%**	**6.76%**	
	fraction of children entering school at age 7	*21.16%*	*35.55%*	*66.17%*	*63.43%*	41.18%
	average years spent in childcare	4.089	3.357	2.735	2.005	3.203
April	**frequency**	**4.40%**	**70.41%**	**17.92%**	**7.26%**	
	fraction of children entering school at age 7	*28.39%*	*43.28%*	*72.70%*	*69.53%*	49.26%
	average years spent in childcare	4.122	3.434	2.805	2.052	3.251
May	**frequency**	**3.44%**	**69%**	**20.07%**	**7.48%**	
	fraction of children entering school at age 7	*39.76%*	*56.27%*	*79.90%*	*74.28%*	61.38%
	average years spent in childcare	4.188	3.565	2.881	2.098	3.340
June	**frequency**	**2.19%**	**58.05%**	**29.87%**	**9.89%**	
	fraction of children entering school at age 7	*74.24%*	*90.90%*	*97.02%*	*95.80%*	92.75%
	average years spent in childcare	4.455	3.911	3.040	2.247	3.498
July	**frequency**	**2.02%**	**55.07%**	**32.56%**	**10.34%**	
	fraction of children entering school at age 7	*80.51%*	*95.30%*	*98.35%*	*97.42%*	96.20%
	average years spent in childcare	4.418	3.955	3.044	2.279	3.494
August	**frequency**	**1.84%**	**52.29%**	**35.18%**	**10.69%**	
	fraction of children entering school at age 7	*83.42%*	*97.22%*	*98.91%*	*98.41%*	97.71%
	average years spent in childcare	4.405	3.976	3.049	2.283	3.477

Note: Bold percentages in each row add up to 100%. The averages are shown for given month of birth.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 10) 2008-2017.

Table H2: ***Effect of Quarter of Birth and the Fraction of 3-Years-Olds in Childcare*** **- First-stage Results on Starting School at Age 7 (D_i) and on Years in Childcare ($yearsCC_i$), Administrative Data – Grade 10**

	cutoff x_d : *January 1*		*cutoff* x_d : *June 1*	
	D_i	$yearsCC_i$	D_i	$yearsCC_i$
Panel A: all students				
$1\{X \geq x_d\}$	0.0859***	0.0317**	0.2663***	0.0729***
	[0.008]	[0.015]	[0.009]	[0.016]
$frac_3yo_in_CC$	0.0120**	0.3344***	0.0704***	0.2500***
	[0.006]	[0.017]	[0.010]	[0.016]
$1\{X_i \geq x_d\} \times$	0.0281***	-0.1021***	-0.0844***	0.0200
$frac_3yo_in_CC$	[0.010]	[0.019]	[0.012]	[0.020]
R^2	0.151	0.043	0.341	0.080
N	211,628	211,628	193,885	193,885
joint F-statistic:	*298.81*	*175.02*	*536.82*	*170.72*
Cragg-Donald Wald F-statistic:	*280.03*		*437.61*	
Kleibergen-Paap rk Wald F-statistic:	*217.88*		*130.62*	
Panel B: by gender				
	male *students*			
$1\{X \geq x_d\}$	0.0847***	0.0447**	0.2225***	0.0259
	[0.012]	[0.020]	[0.012]	[0.021]
$frac_3yo_in_CC$	0.008	0.3169***	0.0799***	0.2520***
	[0.009]	[0.021]	[0.013]	[0.022]
$1\{X_i \geq x_d\} \times$	0.0583***	-0.0845***	-0.0961***	0.0076
$frac_3yo_in_CC$	[0.014]	[0.025]	[0.015]	[0.026]
R^2	0.164	0.038	0.299	0.065
N	102,691	102,691	94,466	94,466
joint F-statistic:	*194.92*	*96.51*	*212.06*	*79.56*
Cragg-Donald Wald F-statistic:	*181.04*		*191.13*	
Kleibergen-Paap rk Wald F-statistic:	*124.00*		*78.40*	
	female *students*			
$1\{X \geq x_d\}$	0.0871***	0.0191	0.3114***	0.1214***
	[0.010]	[0.019]	[0.012]	[0.021]
$frac_3yo_in_CC$	0.0156**	0.3505***	0.0624***	0.2496***
	[0.006]	[0.021]	[0.013]	[0.020]
$1\{X_i \geq x_d\} \times$	-0.0005	-0.1179***	-0.0758***	0.0286
$frac_3yo_in_CC$	[0.012]	[0.025]	[0.014]	[0.027]
R^2	0.123	0.052	0.370	0.101
N	108,937	108,937	99,419	99,419
joint F-statistic:	*122.92*	*147.32*	*447.02*	*167.87*
Cragg-Donald Wald F-statistic:	*104.01*		*251.50*	
Kleibergen-Paap rk Wald F-statistic:	*107.38*		*102.87*	

Notes: [1] The table shows the first-stage results of D_i and CC_i, where D_i is a binary variable denoting age at primary school entry of child i (1: child entered primary school at age 7, 0: child entered primary school at age 6), and CC_i is a binary variable denoting age at childcare entry of child i (1: child entered childcare at age 3+, 0: child entered childcare at age 2). $1\{X \geq x_d\}$ is a binary variable denoting quarter of birth of child i (1: child was born on/after January 1, 0: child was born before January 1). $frac_3yr_in_childcare_i$ denotes the fraction of 3-year-old children in childcare in child i's municipality at her age of 2. Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls.
[2] The sample includes children born in the three-months window around the cutoff dates of January 1.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 6) 2008-2017, KIR-Stat/DEM 1998-2006.

Table H3: **Second-stage//IV Results of Starting School at Age 7 (D) and Years in Childcare ($yearsCC$)**

Panel A: outcome Y in grade 10:	*cutoff x_d : January 1*			*cutoff x_d : June 1*		
	math	**reading**	$1\{repeat\}$	**math**	**reading**	$1\{repeat\}$
	***male** students*					
$D = 1\{$**start school at 7**$\}$	0.2621**	0.2746***	-0.0740**	0.0177	0.0748	-0.0372
	[0.131]	[0.094]	[0.030]	[0.084]	[0.078]	[0.026]
$yearsCC$	0.2410***	0.1455*	-0.0407*	0.2684***	0.2283**	-0.0461*
	[0.069]	[0.083]	[0.022]	[0.097]	[0.092]	[0.024]
N	102,637	102,647	102,691	94,409	94,431	94,466
Hansen J-statistics:	*0.899*	*2.104*	*0.562*	*0.101*	*1.31*	*0.512*
	***female** students*					
$D = 1\{$**start school at 7**$\}$	0.1884*	0.0229	0.0487	-0.1110*	-0.1017*	-0.0524***
	[0.096]	[0.130]	[0.038]	[0.064]	[0.061]	[0.017]
$yearsCC$	0.2127**	0.1626**	0.0141	0.3224***	0.2984***	0.0455**
	[0.086]	[0.064]	[0.016]	[0.087]	[0.083]	[0.019]
N	108,859	108,921	108,937	99,352	99,404	99,419
Hansen J-statistics:	*0.379*	*1.815*	*0.588*	*0.026*	*0.149*	*0.136*
Panel B: HS tracks in grade 10:	*cutoff x_d : January 1*			*cutoff x_d : June 1*		
	low	*middle*	*high*	*low*	*middle*	*high*
	***male** students*					
$D = 1\{$**start school at 7**$\}$	-0.1499***	0.0147	0.1352***	-0.0750**	-0.0078	0.0828**
	[0.038]	[0.051]	[0.045]	[0.032]	[0.043]	[0.041]
$yearsCC$	0.0076	-0.1076*	0.1	-0.0253	-0.0832	0.1085
	[0.042]	[0.062]	[0.072]	[0.046]	[0.071]	[0.082]
N	102,691	102,691	102,691	94,466	94,466	94,466
Hansen J-statistics:	*0.447*	*0.581*	*0.056*	*0.071*	*0.075*	*0.298*
	***female** students*					
$D = 1\{$**start school at 7**$\}$	-0.05	-0.0505	0.1005	-0.0217	0.0079	0.0139
	[0.048]	[0.081]	[0.083]	[0.022]	[0.049]	[0.052]
$yearsCC$	-0.0092	-0.0832	0.0924	-0.0133	-0.0991	0.1124
	[0.027]	[0.064]	[0.071]	[0.029]	[0.083]	[0.090]
N	108,937	108,937	108,937	99,419	99,419	99,419
Hansen J-statistics:	*1.579*	*0.716*	*0.021*	*0.451*	*0.243*	*0.963*
Panel C: educ. aspirations:	*cutoff x_d : January 1*			*cutoff x_d : June 1*		
	$< HS$	HS	$HS+$	$< HS$	HS	$HS+$
	***male** students*					
$D = 1\{$**start school at 7**$\}$	-0.0587*	-0.073	0.1318***	-0.0580**	0.0257	0.0323
	[0.032]	[0.051]	[0.045]	[0.028]	[0.044]	[0.042]
$yearsCC$	0.019	-0.0688*	0.0498	0.0007	-0.0724*	0.0717
	[0.024]	[0.036]	[0.039]	[0.028]	[0.042]	[0.045]
N	100,918	100,918	100,918	92,888	92,888	92,888
Hansen J-statistics:	*0.106*	*0.016*	*0.012*	*0.41*	*0.668*	*0.145*
	***female** students*					
$D = 1\{$**start school at 7**$\}$	-0.0127	0.0357	-0.023	-0.0275*	-0.0228	0.0503
	[0.035]	[0.067]	[0.064]	[0.016]	[0.033]	[0.032]
$yearsCC$	-0.0069	-0.0235	0.0304	-0.008	-0.0543	0.0623
	[0.017]	[0.033]	[0.033]	[0.019]	[0.041]	[0.041]
N	107,475	107,475	107,475	98,066	98,066	98,066
Hansen J-statistics:	*2.61*	*0.097*	*0.443*	*0.007*	*2.055*	*2.338*

I Robustness Check: Accounting for Relative Age Effects

I now investigate whether the positive impacts of redshirting on boys' test scores are robust to accounting for within-class relative age effects. In what follows, I horse-race (absolute) school-starting age D_i and (relative) age rank in class RR_i by instrumenting with birth month and the share of summer-born children in class. Dominating absolute age effects would suggest that redshirting matters, because it boosts boys' human capital, rather than simply making them older relative to their classmates.

First, I construct each child's relative rank (RR) within the class, using date of birth and unique school and class identifiers for all students. In contrast to the data in the reviewed literature, the Hungarian administrative data used in this paper has information about all students' birth month and school-starting age in a given class.[39] Thus, the child's relative age rank within their class is known. Thus, crucially for this data creation step, there is no sample selection based on test score availability or childcare attendance. To create RR, I have data on more than 95 percent of all children. RR is measured in percentiles, where the oldest child in the class receives a value of 100, and children born in the same month have the same value. To instrument for RR, I construct the share of summer-born children in each class ($peers_summer$) similarly.

Then, I augment (2) and also include the child's relative age rank in her class as a regressor:

$$Y_i = \pi_0 + \pi_1 D_i + \pi_2 RR_i + \pi_3 X_i + \pi_4 C_i + \sum_{t=1}^{T-1} \tau_{1t} F_t + \varsigma_i, \tag{4}$$

where RR_i is the child's relative age rank in her class (in percent), and all other variables are defined as before; notably, D_i takes the value of 1 if the child started school at the age of 7, as opposed to 6. Out of the coefficients of interests, π_1 represents the absolute age effect, i.e. the impact of entering school a year older, by comparing two similar children who also have the same position in the age distribution in their class; π_2 represents the relative age effect, i.e., it shows the impact of moving up in the class age distribution by 1 percentile, by comparing two very similar children from the same cohort who started school at the same age.

First, one might wonder where variation in within-class RR comes from, conditional on D (or, in general, absolute age in months), and how large this variation may be. The variation stems from the fact that classes are relatively small, and children in different classes face different age distributions of their classmates (i.e., there is natural variation in classmates' age). Holding the variance of the class age distribution fixed, a student with a given absolute age will have a lower RR in a class with a higher mean absolute age. However, holding the mean, minimum, and maximum absolute age of classes fixed, a student with a given absolute age (and thereby also a constant relative-to-the-class-mean age) will have different ranks (RR) in classes with different variance in absolute age. Note that the same idea is applied by Murphy and Weinhardt (2020), except for test score ranks. Table I3 shows for two select years—2010 and 2014—that there is still substantial variation in within-class RR, even for children of the same age and school-starting age in Grade 6.

In (4), both D_i and RR_i may be endogenous. Besides $1\{X \geq x_d\}$, the two additional instrumental variables are the share of summer-born children in the class and its interaction with $1\{X \geq x_d\}$. The intuition is as follows: Summer-born children start school at age 7 by law and, in the absence of many redshirted peers, are typically the oldest in the class. That is, having a larger share of summer-born children decreases the relative age rank of any summer-born child within their class, *ceteris paribus*.

The share of summer-born children is a valid instrument if it is only related to student achievement (say, test score) through the endogenous variables, D and RR. As opposed to the child's relative age rank, the propensity to start school at age 7 is less likely to be related to the share of summer-born children. Nevertheless, this share violates the exclusion restriction if schools sort children non-randomly into classes based on student achievement. Since summer-born children generally have higher test scores, classes with more high-achieving students would very likely have an over-representation of summer-born children. In order to identify absolute and relative age effects, and to respect the exclusion restriction, it is essential to restrict the sample to children who study in schools that do not sort children systematically into classes based on student achievement.

To separate non-sorting schools from sorting schools, I follow a slightly modified method of Horvath (2015). For each school, I test whether classroom assignment is related to student achievement. I run an *F-test* on the joint significance of classroom effects within each school on mathematics test score (with year fixed effects). The H_0 of this test corresponds to "no sorting," while the H_a corresponds to "sorting."[40] I classify "sorting" schools as those with *p-values* of less than 0.15, a more conservative requirement than the usual 0.05 level.

[39] The class identifiers do not have a meaning across years, and thus cannot be used for estimating class fixed effects.

[40] The intuition behind these tests is that they relate the within-class sum-of-squares of test scores to the between-class sum-of-squares. A sufficiently large within-class variation is indicative of heterogeneous student composition within

Table I1 shows the results from first-stage regressions on the endogenous D and RR variables. Being born between January–March, as opposed to October–December, significantly increases children's likelihood of starting school at an older absolute age of 7, but not their relative age rank in the class. The share of summer-born peers in the class is not strongly related to SSA, as expected, but is significantly related to (relative) class age rank. For fall-born boys (girls), a 1 percentage point increase in the share of summer-born classmates leads to a 0.27 (0.33) percentage point decrease in RR, *ceteris paribus*. For children born between January–March, a 1 percentage point increase in the share of summer-born classmates leads to a 0.06 percentage point increase in RR. Regarding the strength of the IVs, the *joint F-statistic* is 85-100 for D and 140-210 for RR, and the values of both of the *Cragg-Donald Wald F-statistics* and the *Kleibergen-Papp F-statistics* are higher than 80. Thus, weak identification is not a concern (Stock et al. (2002), Olea and Pflueger (2013), Lee et al. (2022)).

Table I1: ***Effect of Quarter of Birth and the Share of Summer-Born Classmates* - First-Stage Results on Higher SSA ($D = 1\{$start school at 7$\}$) and RR, Administrative Data – Grade 6, By Gender**

cutoff x_d : January 1	boys		girls	
	D	$RR(\%)$	D	$RR(\%)$
$1\{X \geq x_d\}$	0.1285***	2.2311***	0.0948***	-0.5874
	[0.0128]	[0.8596]	[0.0105]	[0.7707]
$peers_summer(\%)$	-0.0001	-0.2679***	-0.0003	-0.3263***
	[0.0003]	[0.0188]	[0.0002]	[0.0162]
$1\{X \geq x_d\} \times$	0.00002	0.2090***	0.0002	0.2514***
$peers_summer(\%)$	[0.0004]	[0.0290]	[0.0004]	[0.0260]
R^2	0.176	0.225	0.145	0.264
N	58,233	58,233	61,283	61,283
joint F-statistic:	*103.54*	*142.74*	*84.64*	*210.56*
Cragg-Donald Wald F-statistic:	*98.68*		*82.932*	
Kleibergen-Paap rk Wald F-statistic:	*98.44*		*82.031*	

Notes: [1] $1\{X \geq x_d\}$ is a binary variable denoting quarter of birth of child i (1: child was born on/after January 1, 0: child was born before January 1). $peers_summer_i$ denotes the fraction of summer-born children in child i's class (measured in percentages). Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls.
[2] The sample includes children born in the three-months window around the cutoff dates of January 1. The sample excludes students who study in a "sorting school" based on prior student achievement, as defined in Section 4.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 6) 2008-2017.

Table I2: ***Effect of Higher SSA Due to Academic Redshirting and Relative Age Rank in Class* - LATE Estimates on 6^{th}-Grade Standardized Test Scores, Administrative Data – Grade 6, By Gender**

cutoff x_d : January 1	boys				girls			
Y – **std. test score:**	**math**		**reading**		**math**		**reading**	
$D = 1\{$start school at 7$\}$		0.4071**		0.3190*		0.081		0.2018
		[0.2068]		[0.1902]		[0.1976]		[0.1854]
$RR(\%)$	0.0044**	-0.0026	0.0046**	-0.001	-0.0002	-0.0016	0.0008	-0.0028
	[0.0022]	[0.0028]	[0.0021]	[0.0026]	[0.0026]	[0.0023]	[0.0025]	[0.0020]
R^2	0.197	0.202	0.227	0.235	0.207	0.207	0.276	0.276
N	58,224	58,224	58,226	58,226	61,258	61,258	61,279	61,279
Hansen J-statistic:		*3.873*		*0.6440*		*2.9510*		*0.3380*

Notes: [1] Control variables include: linear trend in month of birth and its interaction with a binary variable denoting quarter of birth, cohort fixed effects, and family background variables listed in Section 4, and missing response characteristics controls.
[2] The sample includes children born in the three-months window around the cutoff dates of January 1. The sample excludes students who study in a "sorting school" based on prior student achievement, as defined in Section 7.
[3] Standard errors are in brackets, and are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.
Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 6) 2008-2017.

classes, which implies classes are roughly similar to each other, so that "no sorting" cannot be rejected.

Table I2 shows the second-stage / IV results for the effect of D and RR on Grade 6 test scores,[41] and reveals that RR has no additional causal effect on test scores above and beyond D. Boys who enter primary school at age 7 are predicted to have higher test scores than children who enter at age 6 by 0.32-0.4 SDs and have the same ranking in the class' age distribution. The values of the *Hansen J-statistic* are below 5, thus the null hypothesis of valid (excludable) instruments cannot be rejected, even at the 1% significance level. The result implies that absolute age effects dominate within-class relative age effects for boys. This result is consistent with Cascio and Schanzenbach (2016), which finds that holding absolute age constant, children who were relatively young in their class performed no worse in school, and with Cook and Kang (2020), which finds that children's absolute age, rather than age relative to classmates, plays the dominant role in parents' decision to redshirt.[42]

Table I3: **Distribution of Relative Rank Variable (RR), by Age in Months and School-Starting Age (SSA), in 2010 and 2014, Administrative Data – Grade 6**

	2010						**2014**					
	$SSA = 7$			$SSA = 6$			$SSA = 7$			$SSA = 6$		
age	***min***	***max***	***sd***	***min***	***max***	***sd***	***min***	***max***	***sd***	***min***	***max***	***sd***
11y9m				0.70	10.00	1.89				0.51	14.29	2.37
11y10m				0.40	25.00	2.79				1.03	23.08	3.00
11y11m				0.47	25.00	3.36				0.36	50.00	3.89
12y0m				0.51	36.36	3.99				0.63	38.46	3.88
12y1m				1.11	52.94	5.50				0.98	40.00	5.68
12y2m				1.39	47.06	6.74				1.56	80.00	6.99
12y3m				1.09	62.50	8.01				1.89	100	7.93
12y4m				2.38	75.00	9.15				3.03	100	9.07
12y5m				3.33	73.08	10.24				3.23	69.23	10.04
12y6m				3.85	100	11.03				4.00	80.00	10.87
12y7m				5.88	81.82	11.38				4.55	82.35	11.49
12y8m				5.88	92.86	11.93				4.35	90.00	12.09
12y9m	6.25	100	12.41	19.35	77.78	17.34	5.56	100	12.67	30.95	66.67	10.01
12y10m	7.14	100	12.55	15.00	88.00	16.07	5.56	100	12.74	30.77	82.14	12.44
12y11m	6.67	100	12.42	37.50	93.33	14.45	9.09	100	12.76	42.86	90.91	12.09
13y0m	12.50	100	12.43	22.73	94.44	14.32	7.14	100	12.26	18.75	100	14.12
13y1m	16.67	100	12.02				5.26	100	11.76			
13y2m	16.67	100	11.16				10.53	100	11.74			
13y3m	14.29	100	11.66				15.38	100	11.16			
13y4m	8.33	100	11.37				15.79	100	10.79			
13y5m	10.00	100	11.82				33.33	100	10.45			
13y6m	27.78	100	11.50				20.00	100	10.98			
13y7m	37.50	100	10.27				21.05	100	10.43			
13y8m	27.27	100	11.22				28.57	100	9.98			
13y9m	7.69	100	13.07				20.00	100	11.79			
13y10m	14.29	100	12.08				38.89	100	9.14			
13y11m	50.00	100	9.99				13.33	100	11.02			
14y0m	42.86	100	10.53				28.57	100	11.85			

Source of data: Administrative test score data – Hungarian National Assessment of Basic Competences (grade 6) 2008-2017.

[41] Results for higher grades and other student achievement outcomes are similar, and are available upon request.

[42] My results are less comparable with that of Bedard and Dhuey (2006), Dhuey and Lipscomb (2008), and Dhuey and Lipscomb (2010), as they do not horse-race school-starting age and relative age rank in the same way that I do.