Essays in Economics of Education

By

ESTELLE (HYEWON) SHIN

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Scott Carrell, Chair

Marianne Bitler

Jenna Stearns

Committee in Charge

2025

Essays in Economics of Education

Abstract

This dissertation examines how educational environments shape students' human capital, particularly social-emotional skills and mental health. The first chapter investigates the persistence of old-for-grade effects on social-emotional skills. The second chapter studies the effects of single-sex schools on social-emotional skills and academic achievement. The third chapter examines how changes to the school-entry cutoff date affect students' mental health.

The first chapter exploits a fuzzy regression discontinuity design and the unique contexts of Seoul Education Longitudinal Study 2010 (SELS 2010). Outcomes include self-esteem, friendships, learning approaches, and goal-setting mindsets, tracked from middle through high school. I find that older girls within a grade exhibit persistently higher self-esteem through high school, while boys show no significant differences. The results also suggest that old-for-grade girls maintain closer friendships and more effective learning approaches. To separate developmental age differences—a potential channel of fadeout—I reconstruct outcome variables to measure skills at the same age. The findings indicate that relative age, rather than developmental age differences, is the primary driver of the old-for-grade effects.

The second chapter exploits unique random middle school assignments in Seoul to examine the effects of single-sex schools on students' social-emotional skills and academic achievement. To isolate middle school effects from those of elementary school and to account for potential sorting into elementary schools, I control for sixth-grade baseline measures and elementary school fixed effects. The results show that boys in all-boys schools exhibit improved self-esteem and learning approaches compared to their counterparts in coeducational schools. Additionally, all-boys schools raise boys' test scores in math and English. Conversely, girls in all-girls schools do not consistently experience improvements in social-emotional skills, although there is suggestive evidence of higher test scores compared to the girls in coeducational schools.

The third chapter, co-authored with Baiyu Zhou, employs a difference-in-difference design and detailed mental health survey data from the Korea Children and Youth Panel Survey 2010 and 2018. Our results suggest that following the reform, both boys and girls born in January and February showed substantially improved mental health during early adolescence, compared to their counterparts born in the same months who entered the school before the policy change. Specifically, girls exhibit lower levels of inattention and aggression, while boys show higher self-esteem. Our findings add novel evidence to the growing body of economics research on mental health during childhood and adolescence.

Acknowledgement

Looking back on my PhD journey, I feel immense gratitude for the many people who helped me navigate this long and uncertain road. There were times I truly doubted whether I could finish. What helped me was not my own effort alone, but the steady support of mentors, friends, and family.

I owe my deepest thanks to my advisor, Professor Scott Carrell, for his continued support and advice with regular check-in. He believed in me even when I couldn't believe in myself, and that trust gave me the strength to move forward. I was also fortunate to have two exceptional committee members. Professor Marianne Bitler consistently offered encouragement, thoughtful advice, and practical help to support me through the job market. Professor Jenna Stearns provided concise and incisive feedback that strengthened my work.

I also thank Professor Shu Shen for invaluable methodological advice, and Professor Katheryn Russ, whose former advisee's advice helped shape my decisions about the future. I'm grateful as well to everyone who participated in the student talk series and teatime, where their engagement help me grow as a researcher.

Even in the hardest moments of Ph.D. life, I was lucky to be surrounded by friends who always stood by me. Eunju Lee helped me through the isolating moments by working together over Zoom. Baiyu Zhou contributed to my research and well-being–sharing concerns and enjoying everyday life. I also thank Justin Ku, Alexa Ludke, GaYoung Park, Minji Kwak, Nayoung Lee, Kyoung-un Song, Jiyoung Lee, Jiyeon Park, and everyone who helped me stay grounded and connected. To my cohort, thank you for walking this path alongside me.

I am especially grateful to my family who supported my choices without pressure or judgment, even during my planned leave of absence during the pandemic. Their encouragement gave me the space to take care of myself and return when I was ready. This dissertation reflects not only my effort but also the presence and influence of many others. I am sincerely thankful to all who made this possible.

Contents

List of Figures

List of Tables

CHAPTER 1

The Persistent Effect of Being Old for Grade on Social-Emotional Skills

1.1. Introduction

Cognitive skills alone cannot fully explain disparities in labor market outcomes. This has prompted researchers to recognize social-emotional skills as a crucial component of human capital (Heckman et al., 2006; Lindqvist and Vestman, 2011). Social-emotional skills, also known as noncognitive skills, are strong predictors of employment, wages, and occupational choices.[1] Importantly, social-emotional skills remain more malleable during adolescence than cognitive skills (Cunha et al., 2010; Hoeschler et al., 2018).While a growing body of literature in economics has examined the impact of schools (Jackson et al., 2020a), teachers (Gong et al., 2018; Kraft, 2019), and peers (Gong et al., 2021) on developing these skills, critical gaps remain in understanding how age-related dynamics within schools shape the development of social-emotional skills over time.

This study examines the persistence of old-for-grade effects on social-emotional skills from middle through high school. Old-for-grade students start school at an older age and continue to be older than their peers throughout their schooling. At the beginning of school, students who start school at an older age often exhibit better achievement in cognitive, physical, and social-emotional domains compared to their younger peers.[2] However, it is unclear whether

[1]Examples include noncognitive measures combining locus of control and self-esteem (Heckman et al., 2006), social skills (Deming, 2017; Weinberger, 2014), and interpersonal skills (Borghans et al., 2008b).

[2]Students who start at an older age are more ready for kindergarten (Dhuey et al., 2019), perform better on tests (Datar, 2006; Elder and Lubotsky, 2009; Cascio and Schanzenbach, 2016), and show fewer behavioral problems and better pro-social skills (Datar and Gottfried, 2015; Lubotsky and Kaestner, 2016) in early schooling.

these "old-for-grade" effects on social-emotional skills fade out as younger students catch up developmentally or persist, creating lasting disparities. Two potential channels can affect persistence. First, if early advantages stem primarily from developmental age differences, the gaps in social-emotional skills may fade out as younger students reach the same developmental stage as older students in previous grades. Alternatively, relatively older students may continue to benefit from their initial advantages due to dynamic complementarity (Cunha and Heckman, 2007), which could reinforce social-emotional growth throughout schooling. Should this gap persist, younger students may experience disadvantages later on, potentially affecting their educational attainment and subsequent labor market outcomes.

To study this, I leverage the unique educational contexts of Seoul, South Korea, using data from the Seoul Education Longitudinal Study 2010 (SELS 2010). First, grade repetition or grade skipping is rare, which helps avoid confounding old-for-grade effects. Second, random assignment to middle schools in Seoul, which is uncommon in other settings, minimizes biases related to selection into higher-quality secondary schools. Third, the 99.7 percent high school continuation rate, combined with extremely low dropout rates, allows me to analyze longer-term effects near labor market entry. This alleviates the concern of nonrandom attrition in high school. Furthermore, the annual panel dataset, which consistently surveyed the same questions across waves, provides a unique opportunity to measure outcomes both at the same age and the same grade, facilitating the decomposition of developmental and relative age effects.

I employ a fuzzy regression discontinuity design, exploiting variations in birthdates around the school starting age cutoff. Since some parents delayed their children's school entry, I instrument for being old for grade using birthdates after the cutoff. This approach follows previous studies that apply similar identification methods.[3] The fuzzy regression

[3]Elder (2010), Evans et al. (2010), Landersø et al. (2017, 2020), and Johansen (2021).

discontinuity design relies on the standard assumptions of regression discontinuity and instrumental variable designs. Additionally, the monotonicity assumption must hold to interpret the estimates as local average treatment effects (LATE). Monotonicity implies that individuals who are affected by the treatment should be affected in the same direction. In the context of my study, students born around the cutoff are expected to become old for grade if they are affected, a condition that is likely to hold.

I find novel evidence that old-for-grade students exhibit persistently higher social-emotional skills than their younger peers throughout middle and high school, with notable differences by gender. Old-for-grade girls show higher self-esteem, with suggestive evidence of stronger friendships and more effective learning approaches compared to younger girls. The magnitude of these effects remains consistent from middle to high school, suggesting persistence over time. In other words, old-for-grade girls demonstrate better social and interpersonal skills, as well as greater motivation and engagement in academic achievement. In contrast, old-for-grade boys show little difference in social-emotional skills compared to younger boys, except for suggestive evidence of stronger friendships in middle school. These effects are not explained by differences in contemporaneous test scores or physical advantages.

Compared to other studies on similar outcomes, the old-for-grade effects on social-emotional skills are relatively large. These substantial effect sizes may reflect the local nature of estimates from both regression discontinuity and instrumental variable designs. While these effects may not be generalized outside the bandwidths or to non-compliers, I compare the reduced-form estimates with those from other studies. For example, the effects on girls' friendships, which range from 0.683 to 0.964 standard deviations, are similar to the impact of increasing the proportion of female peers in a classroom by 33.5 to 48.5 percentage points (Gong et al., 2021).

To further investigate channels of developmental age differences and relative age compared to grade peers, I reconstruct outcome variables to measure skills at the same age, which is possible only because of the annual panel data. To my knowledge, this is the first study to examine whether developmental age differences drive old-for-grade effects on social-emotional skills using this approach. In these age-based measures, young-for-grade students would be the same age as old-for-grade students but with one more year of education. If being one year younger at the time of the survey caused the skill gaps, or if the additional year of schooling compensates for the old-for-grade advantages, the effect should be close to zero. However, the results show that the estimates remain similar and statistically significant, suggesting that relative age is more likely the driving factor.

This paper makes several contributions to the literature on school starting age. First, I provide clean causal evidence of persistent old-for-grade effects on social-emotional skills. Second, I decompose these effects into developmental and relative age components, which are crucial for identifying the channels through which fadeout and persistence occur. My approach builds on previous studies on health (Johansen, 2021), crime (Landersø et al., 2017), and parental outcomes (Landersø et al., 2020). The other approach is the instrumental variable method, using assigned relative age and school starting age (Bedard and Dhuey, 2006), though it may be problematic due to potential violations of the monotonicity.[4] To the best of my knowledge, this is the first paper to decompose developmental and relative age components in social-emotional skills by comparing age-based and grade-based measures. Finally, I uncover gender differences in the development of social-emotional skills. While boys are generally thought to benefit more from being older in grade, old-for-grade girls experience greater gains in social-emotional skills, and these gains persist through higher grades.

These findings have important implications for policy, particularly in designing interventions to support social-emotional development. While I cannot directly observe students

[4]Peña and Duckworth (2018) and Peña (2020) applied this method to grit, empathy, and decision-coping skills.

after high school, previous studies indicate that enhanced social-emotional skills are crucial for labor market outcomes. For instance, a one standard deviation increase in a combined measure of self-esteem and locus of control is associated with a 4.0 percent increase in wages, while a one standard deviation increase in social skills is linked to a 0.4 percentage point increase in full-time employment and a 2.9 percent increase in wages (Deming, 2017). Additionally, higher levels of learning approaches can enhance post-secondary educational attainment. Given the importance of social-emotional skills, this study underscores the potential need to address the challenges faced by young-for-grade girls when designing programs for social-emotional development.

The remainder of this paper is organized as follows. Section 2 describes the educational system in South Korea and the data I use. Section 3 explains the identification strategy, and Section 4 presents the results and interpretation. Section 5 presents robustness checks. Section 6 concludes.

1.2. Educational System and Data

1.2.1. Educational System in South Korea. Education in South Korea consists of six years of primary school, three years of middle school, and three years of high school. The academic year begins in March, and the school starting age cutoff date was March 1st until 2008.[5] Children become eligible to enter primary school in March following their 6th birthday. Consequently, if children enter school on time, March-born children turn seven years old, while February-born children are six years old at school entry. Although early school entry or delayed entry is possible, early entry is relatively rare, with only about 0.3 - 0.6 percent of children in Seoul entering school early each year. In population-level data in Seoul, 779 out of 124,209 children entered school early in 2004, and 377 out of 111,496

[5]The school starting age cutoff moved from March 1st to January 1st in 2008, and the January 1st cutoff was enforced starting in 2009. Children who turn six years old by January 1st are eligible to enter primary school in March of that year, effectively starting school in the calendar year of their 7th birthday. The sample utilized in this paper entered school when March 1st was the school starting age cutoff.

students entered early in 2007. The share of early entrants remained relatively stable until the school entry cutoff moved to January 1st.[6] On the other hand, delayed school entry is more common, with approximately 5.4 percent in 2004 and about 8.7 percent in 2007 delaying their school entry in Seoul.[7] [8]

Before entering primary school, children in South Korea typically attend either kindergarten or daycare centers. Kindergarten is not entirely free when the children in my sample are aged three to five, and kindergarten and daycare centers did not have national curriculum before 2012. The share of children attending either kindergarten or daycare centers at age five was 76.3 percent in 2004 and 87.9 percent in 2007 in Korea. Most children are enrolled in one of these programs, which indicates that many old-for-grade children have one more year of experience in pre-primary education before school entry. This also indicates that old-for-grade and young-for-grade children experience different education cycles at the same age.[9]

Grade repetition is infrequent in South Korea. In 2010, only a small number of students, 12 in primary school and 45 in middle school, repeated grades.[10] Grade repetition may have differential impacts on the development of social-emotional skills compared to students advancing to higher grades. Additionally, young-for-grade students are more likely to have lower test scores, leading to a higher likelihood of repeating a grade (Dhuey et al., 2019). This suggests that grade repetition may result in differential social-emotional development

[6]Source: Statistical yearbook of Seoul education 2005, 2008

[7]To delay school entry, parents need to file an application for reason such as child development, and the primary school headmaster decide whether to accept the application following the deliberation of the compulsory education management committee.

[8]Nationwide delayed entry rates in the United States are similar to those in Korea. In the United States, delayed entry rates range from 5 percent to 7 percent in recent years (Datar, 2006; Bassok and Reardon, 2013; Snyder and Dillow, 2013). Delayed entry in the United States was more prevalent in the 1970s and 1980s, with an average of 12 percent, but has declined over time (Huang, 2015).

[9]Kindergarten provides care for children from age three until they transition to primary school, while daycare centers provide care for children aged zero to six. These two types of early childhood education and care facilities operate under different laws and government supervision. Among children aged five in South Korea who attend kindergarten or daycare centers, approximately 60 percent attended kindergarten, while the remaining attended daycare centers.

[10]Source: Statistical yearbook of education

trajectories for young-for-grade students. In my setting, the infrequency of grade repetition ensures unbiased estimates.

South Korea mandates six years of primary school and three years of middle school. 99.7 percent of middle school graduates advance to high school with an annual high school dropout rate of only one percent.[10] While the minimum school leaving age can potentially influence the length of schooling based on when students start school (Barua and Lang, 2016), the combination of nine years of compulsory education and a substantial advancement rate to high school reduces concerns related to dropout and its impact on social-emotional development.

In Seoul, middle school assignments are random within 11 school districts. High school assignments are based on student preferences. Most students choose between academic and vocational high schools. Within academic high schools, students submit preferences. Random school assignments in middle schools prevent my estimates from being confounded by school quality effects as old-for-grade students are more likely to outperform young-for-grade students in primary school.[11]

1.2.2. Data. I use the Seoul Educational Longitudinal Survey 2010 (henceforth SELS 2010) for this study. SELS 2010 sampled fourth-, seventh-, and tenth-grade students in 2010 and conducted annual surveys until their high school graduation.[12] I utilize data from the fourth-grade and seventh-grade cohorts to analyze the effects in both their high school years (grades 10 to 12) and middle school years (grades 7 to 9). After dropping students with missing dates of birth, there are 3,834 students in the fourth-grade cohort and 4,515

[11]Although high school assignments in Seoul are based on student preferences, the consistency of the effects observed across both middle and high school suggests that preference-based assignments do not significantly alter the impact of being old for grade. The estimates are also robust to the inclusion of school fixed effects.
[12]SELS 2010 utilized a stratified two-stage cluster sampling approach. In the first stage, schools were randomly selected based on the number of schools within each school district. In the second stage, two classes were randomly chosen within each school.

students in the seventh-grade cohort.[13] To address potential concerns regarding non-random attrition in the panel survey, I investigate whether attrition is correlated to being born after March 1st or being old for grade. However, the correlation between attrition and being born after the cutoff or being old for grade is close to zero, suggesting that attrition is unlikely to significantly bias the results.

One limitation of this data set is that it is not possible to directly observe the school starting age. Instead, I use the observed age in the first survey wave as a proxy for the school starting age. If students have repeated a grade or experienced grade retention, the observed age may not perfectly align with the actual school starting age. However, this is less problematic due to the infrequency of repetition or retention in South Korea. Retention typically occurs when a student's attendance rate falls below two-thirds of the total school days. In 2007, only 9 primary school and 23 middle school students repeated grades in Seoul. Furthermore, students can postpone school enrollment, but this is only allowed for reasons such as illness or other extenuating circumstances.[14] In this sample, the majority of students in the seventh-grade cohort were born in 1997 and 1998, while the majority in the fourth-grade cohort were born in 2000 and 2001. Only a small number of students are older or younger than their assigned school age, which reassures that the observed age in the fourth or seventh grade likely reflects the actual school starting age.

The main outcomes of interest are self-esteem, friendship, learning approaches, and goal-setting mindsets. These social-emotional skills have been demonstrated to have an effect on academic attainment and labor market outcomes in previous studies. For instance, studies

[13]Some students are added in survey waves 6 and 7. 164 students in the fourth-grade cohort and 33 students in the seventh-grade cohort are added for a new policy evaluation in the sixth survey wave. 75 students are additionally added for the fourth-grade cohort in the seventh survey wave.

[14]Elementary and Secondary Education Act Article 28 Paragraph (4): The head of an elementary school or a middle school shall grant the exemption or postponement of the obligation of school enrollment pursuant to paragraph (2) or (3) only where the application thereof is filed because of illness or any extenuating circumstances determined by the superintendent of education.

have shown that self-esteem affects educational attainment, wage, employment, and occupational choices (Heckman et al., 2006; de Araujo and Lagos, 2013). The friendship index measures a student's social and interpersonal skills, which predict later employment and occupational choices (Deming, 2017; Lindqvist and Vestman, 2011; Borghans et al., 2008b). The learning approaches index measures students' learning strategies, effort, and motivations, all of which can positively affect their educational attainment. Finally, the goal-setting mindsets index measures whether students have clear goals, know how to achieve them, and work hard to attain them, indicating a connection to motivation.

The self-esteem index includes five questions from the Rosenberg Self-Esteem scale, which is widely used in Economics research. These questions are: *I feel that I have a number of good qualities. I am able to do things as well as most other people. I feel that I'm a person of worth. I take a positive attitude toward myself. On the whole, I am satisfied with myself.*

The friendship index includes four questions of the following: *I have a friend whom I can trust and talk to. Rather than being alone during breaks or lunchtime, I spend time with my friends. Even if I have arguments with my friends, we make up quickly. I help my friends who need assistance.*[15]

The self-directed learning approach index includes nine questions based on the motivated strategies for learning questionnaire (MSLQ) (Pintrich et al., 1993), which include: *I think about how to connect newly learned information with what I already know. I study important materials by summarizing key points or organizing them in tables or mind maps. I check myself to see if I understand the class materials well. I try my best to fully understand the school materials. I try to stick to the study schedule I planned as much as possible. If there is anything I don't understand while studying or doing assignments, I look it up in books or*

[15]Some of these questions are similar to the interpersonal skill survey questions in Jackson et al. (2020a) and Datar and Gottfried (2015); Lubotsky and Kaestner (2016).

online. I believe I will eventually understand any difficult material. I find studying enjoyable. I can handle my own work well without anyone telling me what to do.[16]

The goal-setting mindset index includes five questions: *I have a clear goal that I want to achieve. I know what I need to do to achieve my goals. I am working hard to achieve my goals. The study I am doing now will help me achieve my future goals. If I achieve my future goals, I believe I can also contribute to society.*

Students respond to each question using a 5-point Likert scale, ranging from 1 being 'strongly disagree' to 5 being 'strongly agree.'[17] A higher score means better social-emotional skills. I present summary statistics for outcome indices by gender and cutoff among students born between January and April in Table 1.1. Across outcomes and grades, girls born in March or April, and thus more likely to be older, have higher average indices than those born in January or February. On the other hand, for boys, being born after March 1st does not consistently result in higher average indices. Between genders, girls have slightly lower self-esteem than boys, but there is no consistent pattern of girls scoring higher or lower than boys in other social-emotional skills.

To better interpret estimates, I standardize each survey response to mean zero and unit variance within each survey wave and sum relevant questions. I average each outcome over middle school years (grades 7 to 9), high school years (grades 10 to 12), across all grades (grades 7 to 12), and across all ages (ages 13 to 17) to reduce noise from the small sample size.[18] All outcomes are then standardized again. I use ages 13 to 17 when averaging across ages because data is unavailable for young-for-grade students at age 18 and old-for-grade

[16]Questions used to measure the learning approach index are similar to academic effort (work hard index) in Jackson et al. (2020a).

[17]Cronbach's α is reported for the reliability of these measures in the fourth survey wave. Most measures score above 0.7. See Table 1.19 for details.

[18]I also estimate the effects on outcomes at a specific grade or a specific age, but these estimates are consistent with those using pooled averages.

students at age 12.[19] Young-for-grade students are in grade 8 at age 13 and grade 12 at age 17, while old-for-grade students are in grade 7 at age 13 and in grade 11 at age 17.

1.3. Empirical Strategy

I use a fuzzy regression discontinuity design to investigate the old-for-grade effects on social-emotional skills. Despite the exogenous nature of dates of birth around the school entry cutoff, the timing of a child's school entry remains endogenous. Parents may choose to delay their child's school entry if they believe the child is not sufficiently ready for school. Therefore, I adopt being born after the school entry cutoff as an instrumental variable for the binary variable old-for-grade, similar to previous studies such as Evans et al. (2010) and Johansen (2021).

Old-for-grade is a binary variable, equal to one if the school starting age is 7 and zero if the school starting age is 6. Figure 1.1 illustrates the share of old-for-grade children by gender.[20] Children born in January and February are more inclined to delay their school entry as their dates of birth approach the cutoff. There is a jump in the share of old-for-grade children on March 1st, with the majority of children born in March and April being old for grade. Consistent with prior research (Fredriksson and Öckert, 2014; Landersø et al., 2017; Cook and Kang, 2020), boys are more likely to delay school entry compared to girls.

I estimate the following equations in the reduced form and in the first stage:

$$Y_{ic} = \beta_1 + \beta_2 cutoff_i + \beta_3 d_i + \beta_4 d_i * cutoff_i + \lambda_c + X'_{ic}\gamma_1 + \epsilon_{ic}$$

$$oldforgrade_{ic} = \alpha_1 + \alpha_2 cutoff_i + \alpha_3 d_i + \alpha_4 d_i * cutoff_i + \delta_c + X'_{ic}\gamma_2 + e_{ic}$$

where Y_{icg} is the outcome of interest, which includes self-esteem, friendship, learning approaches, and goal-setting mindsets in middle school and high school. d_i is the date of

[19]The survey was conducted in July every year, so the age is calculated based on the survey time.
[20]The share of old-for-grade students can be interpreted as old-for-grade within their respective grade cohorts in Seoul.

birth relative to the March 1st cutoff. It is equal to 0 when a child is born on March 1st, 1 if born on March 2nd, and -1 if born on February 28th. $cutoff_i$ is a binary variable equal to one if children are born after March 1st. I also include an interaction between d_i and $cutoff_i$. λ_c and δ_c are the survey cohort fixed effects in the reduced form and the first stage respectively. X_i is predetermined individual characteristics, including gender, firstborn, parents' educational attainment (categorized as less than high school, high school graduates, and more than high school), and indicators for missing covariates. I do not include individual controls in my main specification because the estimates remain robust to the inclusion of controls. The estimate of interest is hence $\hat{\beta}_2/\hat{\alpha}_2$.

I use a triangular kernel, which is optimal for boundary estimation, using local linear regression (Fan and Gijbels, 1996). In the main analysis, I choose 60-day bandwidths to ensure a comparison of the same sample across outcomes. Moreover, I conduct additional analyses using bias-corrected estimates and robust confidence intervals, utilizing bandwidths that minimize the mean squared errors (MSE) following Calonico et al. (2014). The results, shown in Table 1.14, remain robust. To further validate my approach, I investigate the optimal local polynomial order that minimizes asymptotic mean squared errors, following Pei et al. (2022). I find that local linear estimators consistently have lower asymptotic mean squared errors than local quadratic or cubic estimators. Results are shown in Tables 1.6 and 1.7.

I cluster the standard errors at the date of birth level. Additionally, I consider alternative clustering at the school level, given that SELS 2010 sampled students at the school level, and students within the same school might be exposed to common shock. The alternative clustering does not yield significantly different results compared to my main specification. Results are shown in Table 1.8. Notably, standard errors are larger when clustering at the date of birth level. This suggests that clustering at the date of birth level provides more

conservative results. Therefore, in my main results, I present standard errors clustered at the date of birth level for robust and conservative estimates.

The validity of the regression discontinuity design relies on the assumption of continuous running variables and the smoothness of the regression functions for potential treatment and potential outcomes near the cutoff. One possible threat to identification is that parents may manipulate the date of birth around the school entry cutoff (Shigeoka, 2015). To address this, I first check whether the density of the running variable around the cutoff changes smoothly (McCrary, 2008). Figure 1.2 presents the density of births within 60-day bandwidths around the cutoff with two-day bins. The number of births is not significantly different before and after the cutoff. Second, I formally test the manipulation using the method proposed by Cattaneo et al. (2020). Results shown in Figure 1.3 do not reject that the density around the cutoff is smooth. I also do the same analysis separately for girls and boys in the Figures 1.6 and 1.7.[21] Lastly, I re-estimate the effects with the donut regression discontinuity design in the robustness check. I exclude one, four, and seven days around the cutoff. Results are shown in Figures 1.18 and 1.19. While the magnitudes become somewhat noisy, they remain largely consistent.

Another potential source of bias in the density of birth is the possibility of unique birth patterns in two specific school cohorts. If, for instance, time-specific events affect the two school cohorts in my data, my estimates would be biased. To mitigate this concern, I analyze whether residuals, obtained by regressing running variables on day-of-week and holiday fixed effects, are smooth around the cutoff. The estimates, shown in Figure 1.8, are robust to the inclusion of day-of-week fixed effects. Furthermore, Kim (2021) finds that adjustments in birth timing between February and March from 1997 to 2007 are small and not statistically

[21]One possible concern related to my data is that the distribution of births around the cutoff date might differ from that at the population level since I use only a sample of students from these cohorts in my data. While I do not have access to the daily number of births at the population level in Seoul, I find that the share of births by month in my data is similar to that at the population level in Seoul. This evidence assures that the cohorts in my sample are not drastically different from the overall population in terms of birth patterns.

significant, based on administrative birth certificates in Korea.[22] This evidence suggests that the birth patterns in my sample are not atypical compared to other cohorts and the population at large.

Another way to assess the validity of regression discontinuity design is to check whether individual predetermined characteristics are balanced around the cutoff. Figure 1.4 presents the distribution of females, firstborns, and parents' educational attainment for students born between January and April. The bin size is set to two days. Overall, the predetermined characteristics of children appear to be similar before and after the March 1st cutoff, suggesting no significant imbalance in these characteristics. One exception is that there is a slightly greater proportion of girls among January and February-born children compared to those born in March and April. I present the results separately for girls and boys, in addition to the entire sample.

Since I am using the instrument, standard instrumental variable assumptions, including the relevance of the instrumental variable and the exclusion restriction, must also be satisfied. Additionally, the monotonicity assumption is necessary to evaluate the local average treatment effect (LATE). Monotonicity implies that students born around the cutoff, if they are affected, should be affected in the same direction—specifically, they should become old-for-grade at school entry if affected in my setting. Instrument variables used in previous studies such as legal school entry age may violate the monotonicity assumption (Aliprantis, 2012; Barua and Lang, 2016; Fiorini and Stevens, 2021). Because I instrument a binary variable old-for-grade with being born after the cutoff, 'defiers' refer to parents and children who delay school entry when children are born before the cutoff and enter early when children are born after the cutoff. In other words, children will be seven years old when they are born before the cutoff and six years old when they are born after the cutoff. This seems implausible. For example, if children are not ready for school at age six, they may choose

[22]One explanation provided by Kim (2021) is that school starting age enforcement is flexible, so parents can adjust the timing of their child's school entry when the child is eligible to enter school. This gives less incentive for parents to adjust birth timing.

to be old-for-grade regardless of whether they are born before or after the cutoff. On the other hand, if children are already ready for school at age six, they will enter school on time if they are born before the cutoff. They may enter early or on time if they are born after the cutoff. Hence, it is not likely that the monotonicity assumption is violated.

1.4. Results

1.4.1. First Stage. I present the first-stage regression results in Table 1.2. All regressions include cohort fixed effects. Columns 1 and 2 show estimates for the entire sample, Columns 3 and 4 show estimates for girls, and Columns 5 and 6 show estimates for boys. The results show that children born after the cutoff are 32.3 percentage points more likely to be old for grade in their grade cohort. The effects are larger for girls with 37.6 percentage points compared to boys with 27.8 percentage points. Given that 44.6 percent of girls and 55.3 percent of boys born in January and February delay their entry, boys are more likely to be old for grade when they are born before cutoff. The estimates are significant in all regressions and remain consistent with and without individual controls. In the following results, I only present regression results without controls.

1.4.2. Main Results. Table 1.3 presents the main estimation results for social-emotional skills during middle school and high school. Odd columns present the effects during middle school (grades 7 to 9), while even columns present the effects during high school (grades 10 to 12). Panel A illustrates the results for girls, Panel B for boys, and Panel C for the entire sample. Columns 1 and 2 show the effects of being old for grade on self-esteem. Girls who are old for grade are more likely to exhibit higher self-esteem compared to young-for-grade girls in both middle school and high school. Magnitudes are similar in both periods as well, suggesting persistence of the effects. On the other hand, the effect on self-esteem for old-for-grade boys is relatively small and statistically insignificant. Panel C

suggests that students who are old for grade generally exhibit higher self-esteem, and Panels A and B indicate these effects are primarily driven by girls.

The effects on friendships, in comparison to those on self-esteem, are generally less pronounced. On the whole, students who are old for grade tend to establish more positive relationships with their peers when compared to their young-for-grade counterparts within the entire sample. Old-for-grade girls, again, persistently enjoy closer friendships during both middle and high school. In the case of boys, the estimate on friendships is not statistically significant, despite the effect size resembling that observed in girls during middle school. Intriguingly, as students progress to high school, the distinction between old-for-grade and young-for-grade boys becomes notably diminished.

Columns 3 and 4 present the effects on friendships. In comparison to those on self-esteem, the effects are generally less pronounced. Overall, old-for-grade students tend to have better social and interpersonal skills in relationships with their peers compared to their young-for-grade counterparts within the entire sample. Old-for-grade girls persistently enjoy better friendships during both middle and high school. For boys, however, the estimate on friendships is not statistically significant, despite the similar effect size to girls during middle school. Interestingly, as students progress to high school, the difference between old-for-grade and young-for-grade boys fades out. Note that the evidence for the positive effect on friendships should be taken cautiously, given the sensitivity of the estimates to various specifications and bandwidth choices. Nonetheless, the findings imply that old-for-grade girls experience persistently better friendships, while the positive effects on old-for-grade boys tend to diminish.

Columns 5 and 6 present the effects of being old for grade on learning approaches. Girls who are old for grade tend to show more effective study habits, greater engagement in learning, and heightened dedication when compared to younger girls. Again, these effects persist throughout middle and high school. The estimate of 0.916 standard deviations in high school

is slightly bigger than the estimate in middle school. One plausible explanation could be the higher perceived stakes associated with high school, as test scores in this period directly influence college admissions. Consequently, old-for-grade girls might exhibit enhanced self-control and study behaviors.

Panel A of Columns 7 and 8 suggests that old-for-grade girls tend to show a better ability to set goals. This implies that they are more likely to have clear goals and actively strive to achieve them. However, the estimates are somewhat smaller and statistically significant only at the 10 percent level during middle school, eventually losing their statistical significance in high school. In contrast, the effects of learning approaches and goal-setting mindsets for boys are relatively negligible during middle school, with a shift towards a negative direction in high school, albeit without statistical significance. Taken together, old-for-grade boys are not significantly more likely to have better learning approaches or goal-setting mindsets than young-for-grade boys.

In summary, the old-for-grade effects are more pronounced for girls. Old-for-grade girls show higher self-esteem, closer relationships with friends, and more effective learning approaches. Importantly, these findings suggest that the positive effects for old-for-grade girls persist over time rather than fade out as the students advance to higher grades. In contrast, old-for-grade boys do not show noticeable improvements in their social-emotional skill development.

To provide a broader context in terms of magnitudes, I first compare my estimates to other papers that study the effect of school starting age on similar outcomes. The effect size for self-esteem, ranging from 0.590 to 0.762 standard deviations for the entire sample, is larger than the 0.15 standard deviation identified by Datar and Gottfried (2015) for eighth graders. Notably, the impact on learning approaches for old-for-grade girls does not diminish over time, in contrast to the findings of Datar and Gottfried (2015) and Lubotsky and Kaestner (2016). The magnitude of the estimate in high school is comparable to the 0.63

standard deviation observed in Lubotsky and Kaestner (2016) or the 0.701 standard deviation identified in Datar and Gottfried (2015) during the Fall of kindergarten.

Then I compare my estimates with other interventions studying similar outcomes. This difference can potentially be attributed to variations in estimation methodologies and the local nature of both regression discontinuity and instrumental variable approaches. Despite potential limitations in terms of external validity beyond the specified bandwidths, reduced-form estimates could be employed for comparison with other interventions. Estimation results from reduced-form regressions are presented in Table 1.11. For instance, the effect on girls' friendships, ranging from 0.257 to 0.372 standard deviations, is comparable in magnitude to the effect of increasing the proportion of female peers in the classroom by 33.5 to 48.5 percentage points Gong et al. (2021). This increases students' social acclimation and general satisfaction index. Similarly, the effect on the learning approaches index of 0.353 standard deviation in high school aligns with the magnitude associated with an increase in teacher quality by 2 to 3 standard deviations on students' effort in class (Kraft, 2019).

Since there are multiple outcomes across middle school and high school, I also investigate the possibility of an increased number of significant results due to multiple hypothesis testing. I employ Bonferroni and Sidak corrections to adjust the p-values. I first adjust p-values across outcomes for girls. While the impact on friendship and learning approaches is not robust, the difference in self-esteem between old-for-grade girls and young-for-grade girls remains statistically significant. Furthermore, girls' self-esteem remains significant even after implementing corrections across both genders and outcomes. These findings further support that old-for-grade girls are more likely to exhibit higher self-esteem than young-for-grade girls, although the effects on other social-emotional skills might not be as robust.

1.4.3. Mechanisms and Interpretation. A crucial question in studying age effects is whether the observed old-for-grade effects are simply due to age differences within the same grade. When skills are measured at the same point of time, younger students might

naturally exhibit lower skill levels because they have had less time to develop these skills. It is possible that these younger students would achieve the same skill levels if assessed one year later when they are the same age as their older peers.

To investigate this possibility, I re-create the outcome variables based on the age rather than the grade. For example, at age 17, old-for-grade students are typically in grade 11, whereas young-for-grade students are in grade 12. I take grade 11 responses for old-for-grade students and grade 12 responses for young-for-grade students to create outcome variables at the same age. Thus, old-for-grade students are the same age but have one less year of schooling compared to their young-for-grade peers. Estimates using these new outcome measures control age at the survey while also including the potentially negative effects of having one less year of schooling. In contrast, estimates using grade-based outcome measures combine the effects of school starting age, relative age compared to peers, and age at the time of the survey. If age at the time of the survey is the primary driver of these effects, the estimates based on outcomes measured at the same age should be smaller and not statistically significant. Ideally, using social-emotional skill measures in the later stages of life would fully control for years of schooling and differences in educational cycles. Although the data I use do not include information beyond high school, I can compare the estimates pooled across all grades (grades 7 to 12) and all ages (ages 13 to 17) to partially examine the influence of years of schooling.

Table 1.4 presents the results. For girls, the estimates at all ages are slightly attenuated compared to those at all grades. This suggests that age at the time of the survey contributes to some of the estimated effects in middle and high school years, or that having an additional year of schooling partially mitigates the impact of starting school at an older age. The effects for boys are generally consistent across all grades and ages.[23] This indicates that the old-for-grade effects do not completely disappear when young-for-grade students reach the same age

[23]These findings remain consistent when comparing the estimates of skills measured at a specific grade and those measured at a specific age. Results are shown in Figures 1.10 and 1.11.

as old-for-grade students and that an additional year of schooling does not fully compensate for starting school at an older age. Age at school entry and relative age compared to peers play a more significant role in social-emotional skill development.

Next, I investigate the potential mechanisms underlying these effects. First, I examine whether old-for-grade girls achieve higher test scores that could enhance their self-esteem.[24] The results, presented in Table 1.12, show that while the estimates are imprecise, old-for-grade girls perform better on Korean and math tests than their younger peers in middle school, though this advantage does not persist in high school. In contrast, old-for-grade boys consistently outperform their younger peers in English tests throughout middle and high school. This suggests that better test performance in middle school could be one factor contributing to higher self-esteem in girls, but test performance alone does not fully explain the disparities in girls' self-esteem.

Additionally, I compare estimates using test scores based on grade with those based on age. Recognizing that tests at each grade level reflect different educational cycles, I address this concern by utilizing the vertically scaled scores provided in SELS 2010, which align test scores across grade levels on a common vertical scale (Kim, 2015). The results, shown in Figure 1.12, indicate that the estimates are generally similar in size between grade-based and age-based test scores. As both girls and boys age or advance to higher grades, the estimates tend to become close to zero, following a pattern similar to that in Table 1.12.

The lack of significant effects on test scores, despite the positive impact on learning approaches, may seem counterintuitive. One plausible explanation is that the exams in the survey may not carry as high stakes. In fact, Kim (2011) finds that a one-month increase in school starting age increases the likelihood of entering a 4-year college by 0.42 percentage points within two years after high school graduation in Korea. This suggests that, although

[24]SELS 2010 developed Korean, math, and English tests based on the first-semester curriculum of each grade in the survey year.

the direct influence on test scores may not be immediately evident, the effects of being old for grade on learning approaches could still play a role in later life.

Lastly, I explore whether old-for-grade students have physical advantages. Since height and weight were surveyed only in the first wave, I analyzed the height and weight of each cohort for grades 4 and 7. The results, as shown in Table 1.13, indicate that while the estimates for girls' height are positive, they are not statistically significant. The gaps between old- and young-for-grade boys are smaller and also not significant. In the seventh-grade cohort, the estimates are smaller compared to the fourth-grade. While old-for-grade girls may be taller than their younger peers in fourth grade, this difference fades by seventh grade. Additionally, the correlation between height and self-esteem is stronger for boys than for girls in the seventh-grade cohort, indicating that physical advantages are less likely to be the main driver of girls' higher self-esteem.

I use the Korean Children and Youth Panel Survey 2010 (KCYPS 2010) to complement the previous analysis for physical advantages and run a placebo test on age-based and grade-based estimates.[25] Since this dataset only provides birth month and year, not the exact date, I apply an instrumental variable approach and focus on children born between January and April.[26] Figure 1.13 shows that old-for-grade girls are about 10 cm (4 inches) taller than young-for-grade girls in grade 5, significantly taller in grades 6 and 7, but the height difference diminishes by grade 8. Despite the height difference in age 11, the age-based estimates are mostly near zero. Boys' height results are noisier, but age-based estimates still trend closer to zero compared to grade-based estimates in grades 6 to 8. These findings confirm that the slight differences between age-based and grade-based estimates for social-emotional skills are not coincidental. Moreover, height differences are more pronounced in grades 5 to 7 for girls

[25]I use the fourth- and seventh-grade cohorts in KCYPS 2010, who are in the same grades as my main sample from SELS 2010.

[26]Using month of birth instead of exact birth date as the running variable, I instrument a born-after-cutoff indicator on the old-for-grade indicator. The model controls for the month of birth, the interaction between the month of birth and the born-after-cutoff indicator, and cohort fixed effects. Standard errors are clustered at the month-of-birth level.

and grades 6 to 8 for boys. Thus, contemporaneous height differences are again less likely to be the primary factor behind girls' higher self-esteem.

Improved social-emotional skills for old-for-grade girls have crucial implications. Self-esteem has positively correlated with post-secondary education and later labor market outcomes, including employment and earnings (de Araujo and Lagos, 2013). Based on the estimates provided by de Araujo and Lagos (2013), a back-of-the-envelope calculation suggests that a one standard deviation increase in self-esteem could result in an approximately 25.06 percent wage increase for females, through an increase in years of schooling. Friendship, which is a combined index for social and interpersonal skills, can similarly impact employment. According to Deming (2017), a one standard deviation improvement in social skills predicts a 3.0 percentage point rise in full-time employment. Given the 0.964 standard deviation increase in friendship for girls, this could potentially translate into a 2.89 percentage point increase in female employment. While interpreting these findings, it is essential to be cautious since most previous studies are based on US data, and some date back to the 1980s and early 2000s, which may differ from the current South Korean context. Nevertheless, it remains plausible that old-for-grade girls with enhanced social-emotional skills might attain more years of education and experience improved labor market outcomes in the future.

1.5. Robustness Check

In this section, I show the robustness of the results through a series of tests. First, I replicate the analyses with different bandwidths. Specifically, I employ mean squared error (MSE) optimal bandwidths for individual regressions and for the first stage, both computed following Calonico et al. (2014). I compare bias-corrected estimates and robust confidence intervals with the main estimates. Secondly, I explore the sensitivity of the results to different functional forms by estimating the effects using global linear and quadratic regressions, also utilizing various bandwidths (30, 60, and 90 days). Third, I assess whether a donut regression discontinuity approach, which excludes a small number of observations close to the cutoff,

yields different results. Fourth, I examine whether including day-of-week and holiday fixed effects in the specification affects the main results.

To evaluate the sensitivity to bandwidth choices, I re-estimate the effects using MSE optimal bandwidths, following Calonico et al. (2014). I employ two approaches for computing MSE optimal bandwidths. The first minimizes the mean squared errors for fuzzy regression discontinuity design, generating different bandwidths for each regression—the MSE optimal bandwidths for fuzzy regression discontinuity range from 42 days to 78 days. The second uses the MSE optimal bandwidths computed in the first stage. Thus, I use the same bandwidths within each gender. The first-stage MSE optimal bandwidths are 38 days for girls and 40 days for boys. Results are presented in the Figures 1.14 and 1.15. The estimates maintain similar magnitudes across the different bandwidth choices. Estimates also remain statistically significant except for girls' friendships in high school and learning approaches in all grades. This indicates that the main results are robust and not sensitive to bandwidth choices.

Concerns might arise due to dates of birth being discrete variables, raising questions about the applicability of local linear regression discontinuity design. To address this concern, I replicate the main results using a parametric regression discontinuity design and cluster the standard errors at the date of birth level, following the approach proposed by Lee and Card (2008). The results obtained through global linear polynomials are largely consistent with my main specifications. To further check sensitivity to functional form, I compare estimates using linear and quadratic control functions and different bandwidth choices (30 days and 90 days). Results are shown in the Figures 1.16 and 1.17. Some estimates become noisy when using a smaller bandwidth of 30 days. This might be attributed to a smaller sample size due to narrower bandwidths. Although some estimates lose statistical significance, the magnitudes are consistent in general. Importantly, results for girls' self-esteem remain large and significant across different specifications. This reassures that the effects on girls' self-esteem are robust to changes in functional forms.

Next, I redo the same analyses using a donut regression discontinuity approach. Although manipulation testing does not reject the smoothness of the density of births, one might be worried about the randomness of births around the cutoff, especially considering March 1st is a holiday in Korea. To examine the robustness of the results, I exclude 1 day, 4 days, and 7 days before and after March 1st. Figures 1.18 and 1.19 indicate that the coefficients are robust to the donut approach. While the effects on girls' friendships and goal-setting mindsets become larger and statistically significant when excluding 7 days before/after the cutoff, the magnitudes of other outcomes generally remain similar.

Additionally, I investigate whether the results are sensitive to including day-of-week and holiday fixed effects. Results are shown in Tables 1.14 to 1.17. Columns 1 and 5 show estimates from my main specifications, and Columns 4 and 8 show estimates with day-of-week fixed effects. The estimates remain nearly unchanged across these different specifications. The results from the donut approach and the inclusion of day-of-week fixed effects help address concerns regarding potential birth manipulation around the cutoff.

1.6. Conclusion

In this study, I examine whether the effects of being old for grade on students' social-emotional skills fade out in late adolescence. By employing a fuzzy regression discontinuity design with a binary treatment of being old for grade and a binary assignment based on birth dates relative to the cutoff, I mitigate concerns about potential violations of the monotonicity assumption. Additionally, I leverage the unique educational context of Seoul, characterized by random middle school assignments, infrequent grade repetition and skipping, and a high rate of high school advancement, to control for confounding factors related to school sorting, repetition, and varying lengths of schooling.

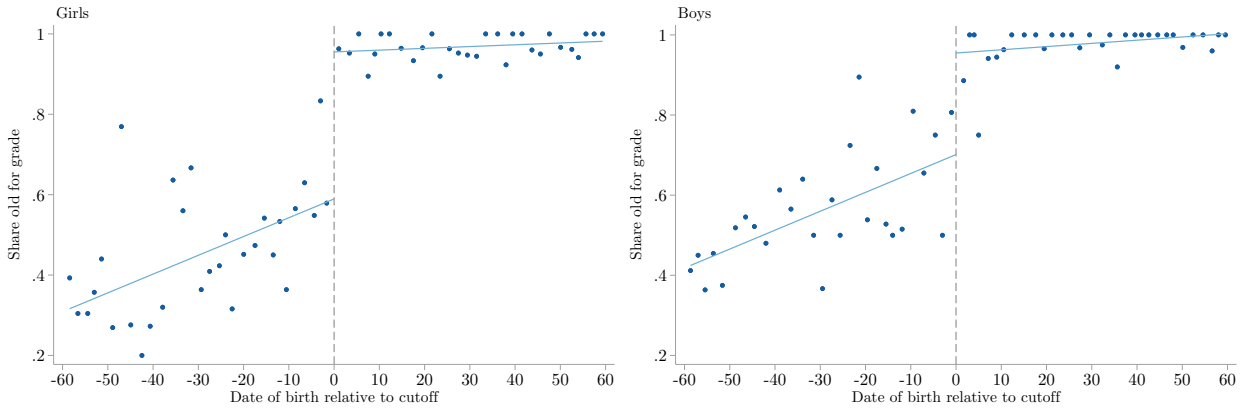
I find that old-for-grade girls consistently exhibit higher self-esteem than their younger counterparts throughout both middle and high school. There is also suggestive evidence that they maintain closer friendships during adolescence and develop better learning approaches in

high school. In contrast, the effects on boys' social-emotional skills are more limited. While old-for-grade boys experience slightly improved friendships in middle school, this effect fades out by high school. These results underscore the distinct trajectories of social-emotional skill development between girls and boys.

Lastly, I distinguish between relative and absolute age effects by comparing estimates from outcomes measured at the same age with those measured at the same grade. Even when controlling for absolute age at the time of the survey, old-for-grade girls continue to demonstrate higher social-emotional skills despite having one less year of schooling. This suggests that relative age is likely the more significant factor driving the old-for-grade effects on social-emotional skills.

This research has important implications for educators and policymakers, emphasizing the in-school program designs that support the development of social-emotional skills. In a school system with a defined cutoff date, there will always be age differences within the same grade, and younger girls may be at a disadvantage in terms of social-emotional skills. The persistent old-for-grade effects into late adolescence suggest that these disadvantages could have long-term consequences for young-for-grade students. Moreover, if, as the theory of dynamic complementarity suggests, skill begets skill, then old-for-grade students may derive greater benefits from interventions aimed at enhancing social-emotional skills. Therefore, it is crucial to consider these gaps when designing in-school programs to support skill development.

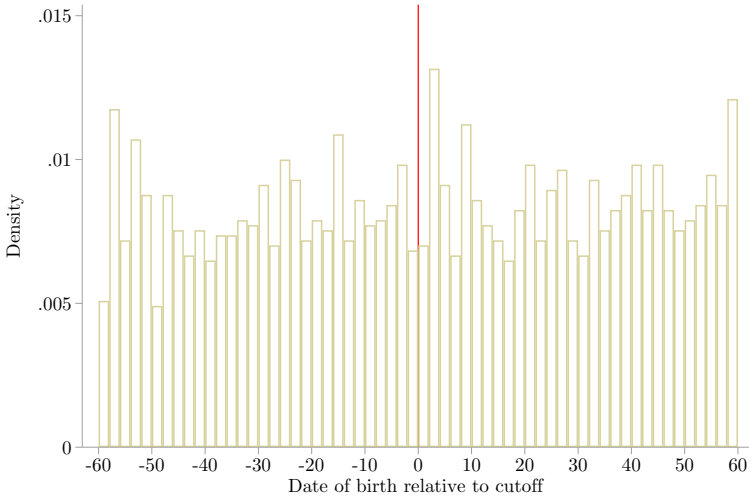
1.7. Figures



(a) Girls (b) Boys

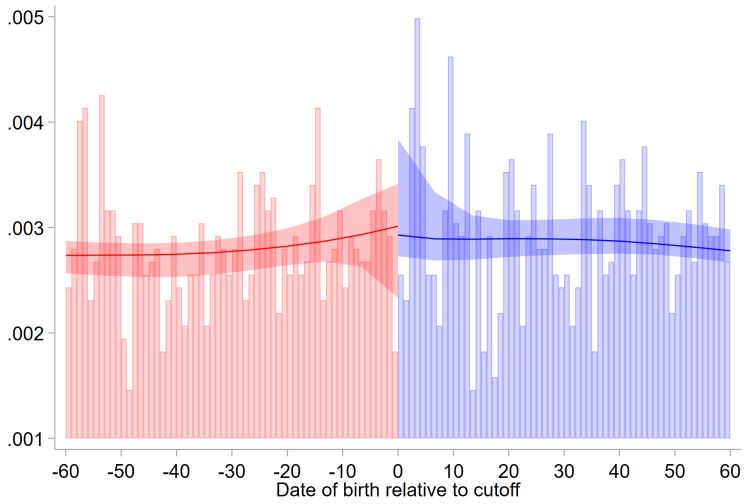
The share of old-for-grade students born between January and April is shown. The sample I use includes fourth-grade and seventh-grade cohorts in SELS 2010. The bin size is set to two days.

FIGURE 1.1. Share of Old-for-Grade Students among Girls (Left) and Boys (Right)



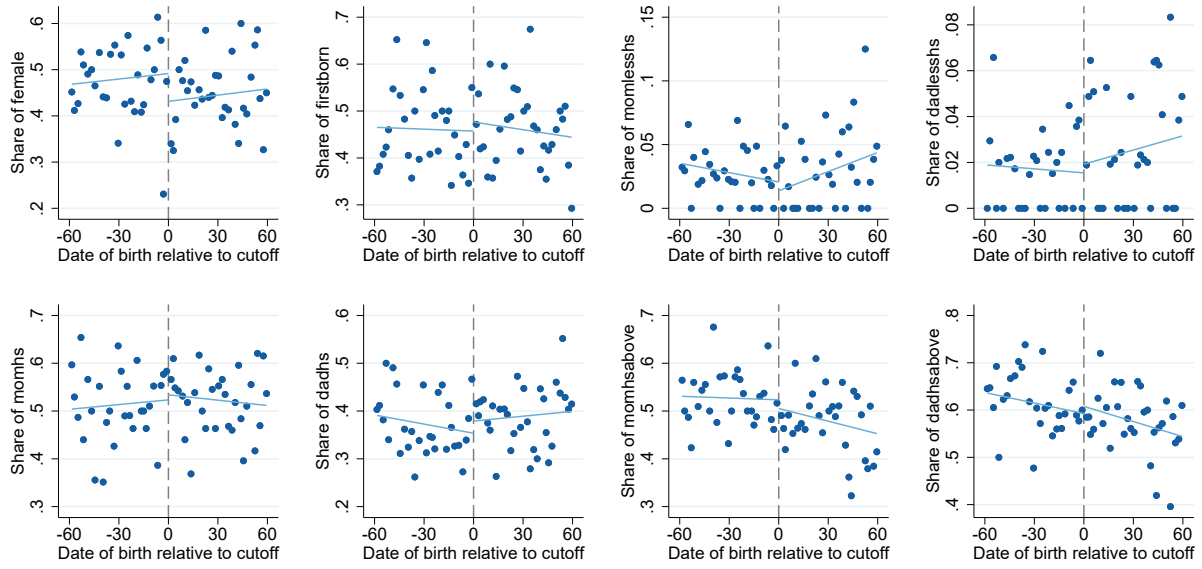
The density of births born between January and April is shown. The sample I use includes fourth-grade and seventh-grade cohorts in SELS 2010. The bin size is set to two days.

FIGURE 1.2. Density of Births Born Around the Cutoff



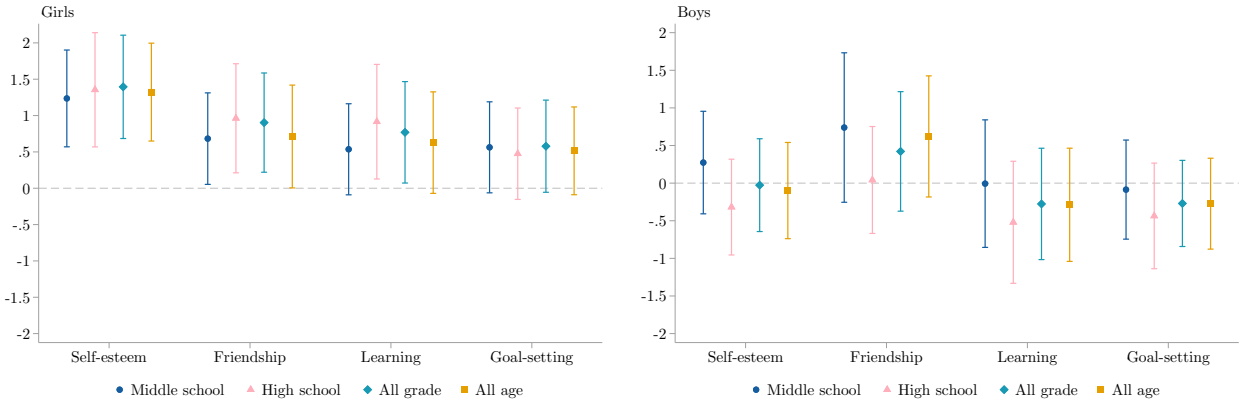
The figure presents manipulation testing of birth density around the cutoff following Cattaneo et al. (2020). The bandwidth is 60 days and the bin size is set to one day. The sample I use includes fourth-grade and seventh-grade cohorts in SELS 2010.

FIGURE 1.3. Manipulation Testing of Births



Each figure presents individual predetermined characteristics within 60-day bandwidths. Each dot represents the mean of each characteristic in a two-day bin. Predetermined characteristics include female, firstborn status, and parents' educational attainment. Parents' educational attainment is divided into three categories: less than high school, high school graduates, and some college or more. I also check for indicators of missing values in these characteristics, which are also balanced around the cutoff.

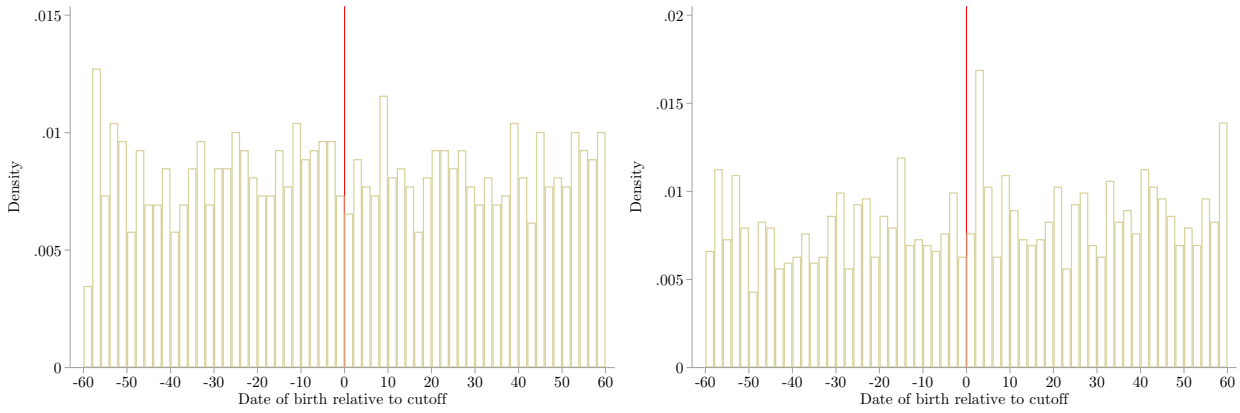
FIGURE 1.4. Comparison of Predetermined Characteristics Before and After the Cutoff



(a) Girls (b) Boys

All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. Middle school (MS) outcomes are averaged across grades 7 to 9, high school (HS) outcomes are averaged across grades 10 to 12, and all grade outcomes are averaged across grades 7 to 12. All age outcomes are averaged across ages 13 to 17. All outcomes are standardized with a mean of zero and a unit variance.

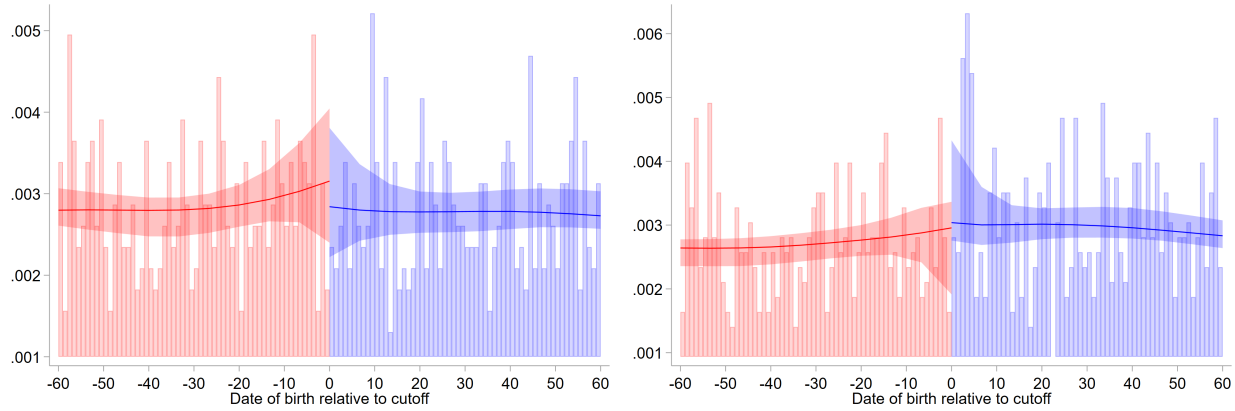
FIGURE 1.5. Effects of Being Old for Grade on Girls' and Boys' Social-Emotional Skills



(a) Girls (b) Boys

The figure shows the density of births born between January and April by gender. The sample I use includes girls and boys in fourth-grade and seventh-grade cohorts in SELS 2010. The bin size is set to two days.

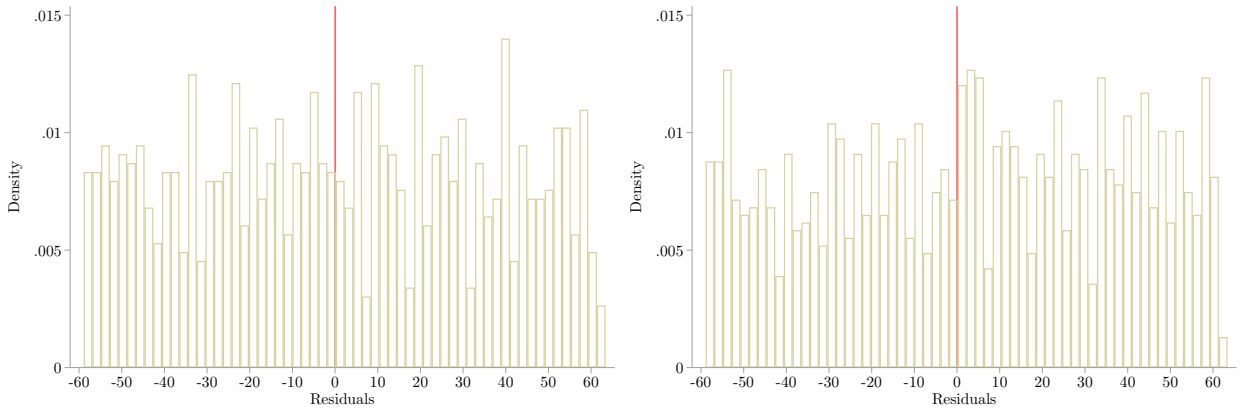
FIGURE 1.6. Density of Births Born Around the Cutoff by Gender



(a) Girls (b) Boys

The figure presents manipulation testing of birth density around the cutoff by gender following Cattaneo et al. (2020). The bandwidth is 60 days and the bin size is set to one day. The sample I use includes girls and boys fourth-grade and seventh-grade cohorts in SELS 2010.

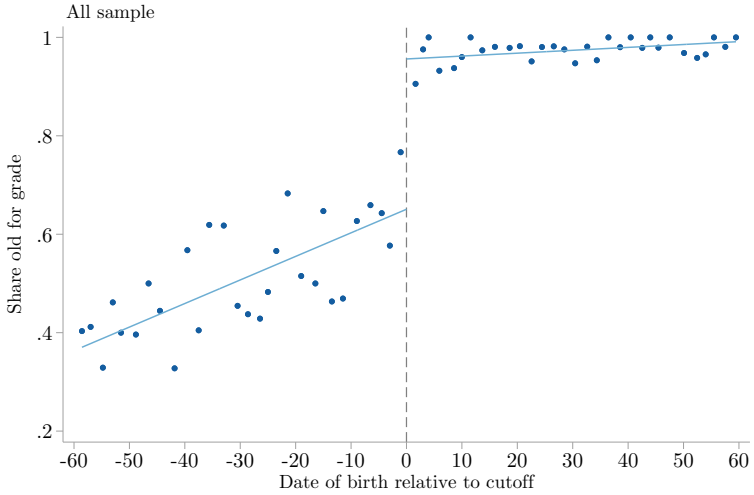
FIGURE 1.7. Manipulation Testing of Birth Density by Gender



(a) Girls (b) Boys

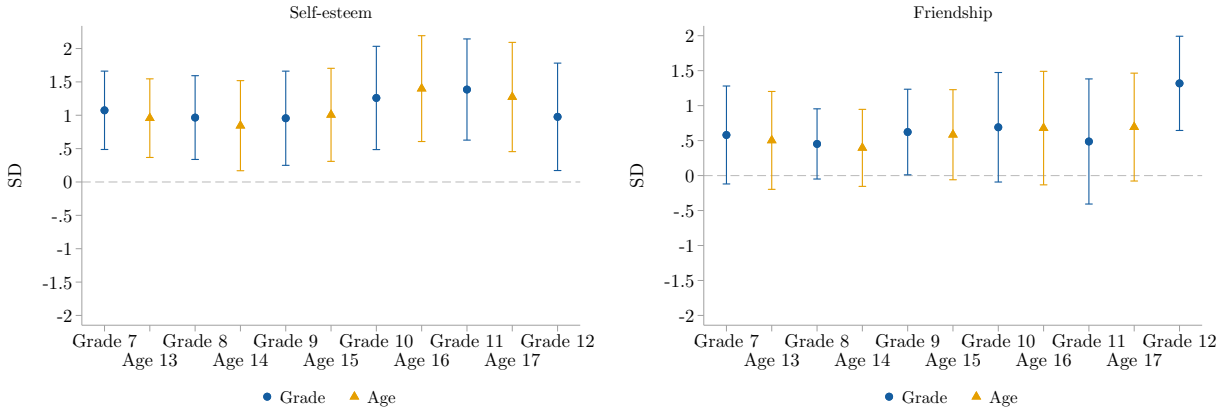
The figure shows the density of births born between January and April, using residuals from regressions conditioning on day-of-week fixed effects and holidays. Panel A shows the birth patterns of girls and Panel B shows the birth patterns of boys. The sample I use includes fourth-grade and seventh-grade cohorts in SELS 2010. The bin size is set to two days.

FIGURE 1.8. Density of Birth Born Around the Cutoff Conditional on Day-of-Week Fixed Effects



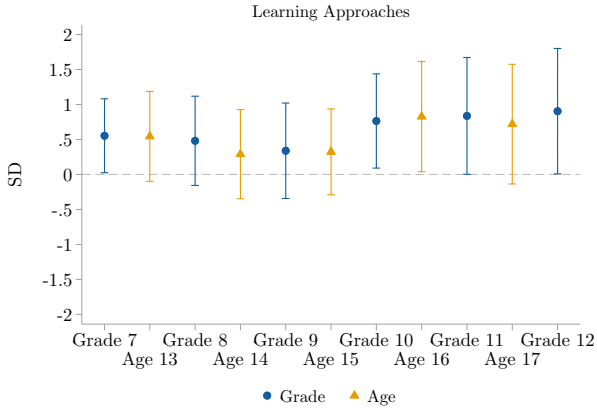
The share of old-for-grade students born between January and April is shown. The sample I use includes fourth-grade and seventh-grade cohorts in SELS 2010. The bin size is set to two days.

FIGURE 1.9. Share of Old-for-Grade Students

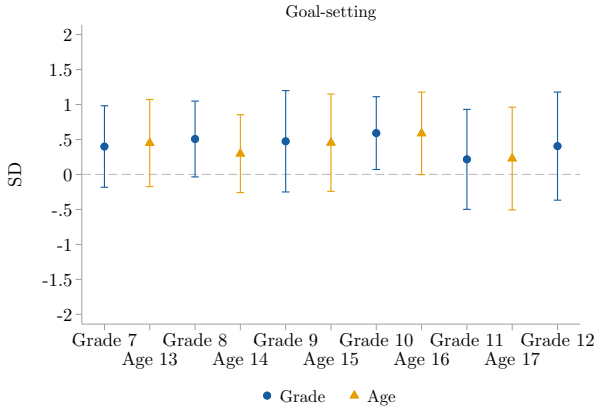


(a) Self-esteem

(b) Friendship



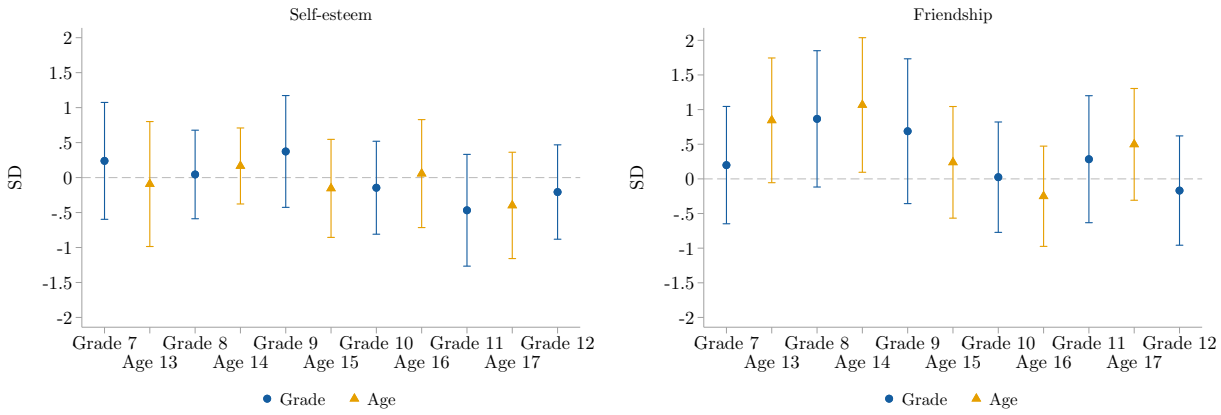
(c) Learning approaches



(d) Goal-setting mindsets

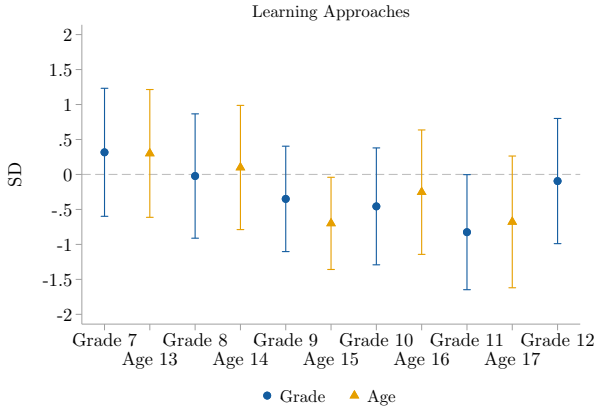
The estimates and corresponding 95 percent confidence intervals (CI) are shown. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.10. Comparison of Old-for-grade Effects in Grades 7–12 and Ages 13–17 (Girls)

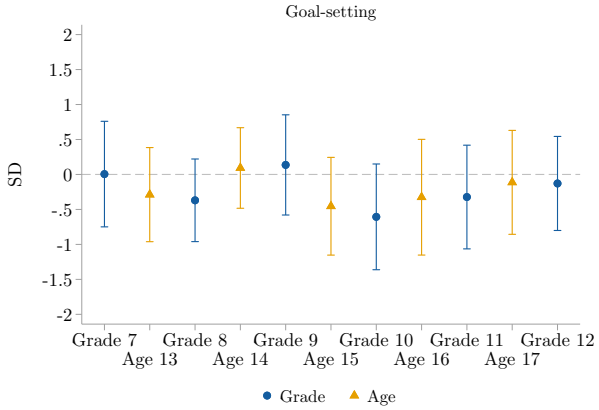


(a) Self-esteem

(b) Friendship



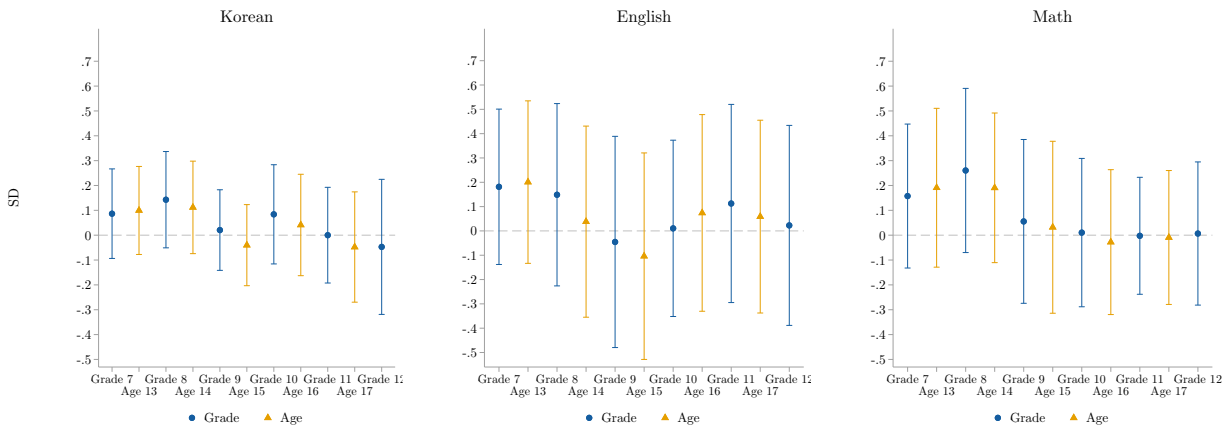
(c) Learning approaches



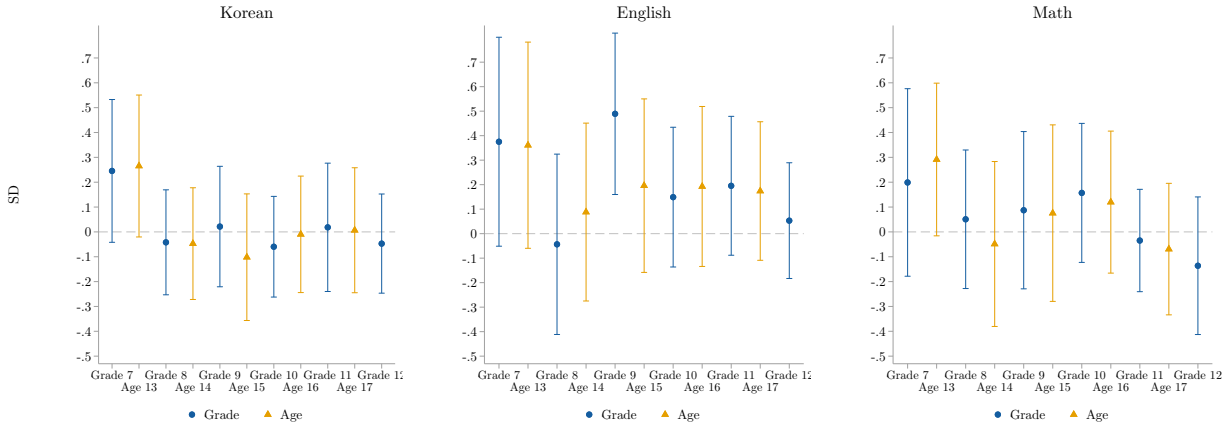
(d) Goal-setting mindsets

The estimates and corresponding 95 percent confidence intervals (CI) are shown. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.11. Comparison of Old-for-grade Effects in Grades 7–12 and Ages 13–17 (Boys)



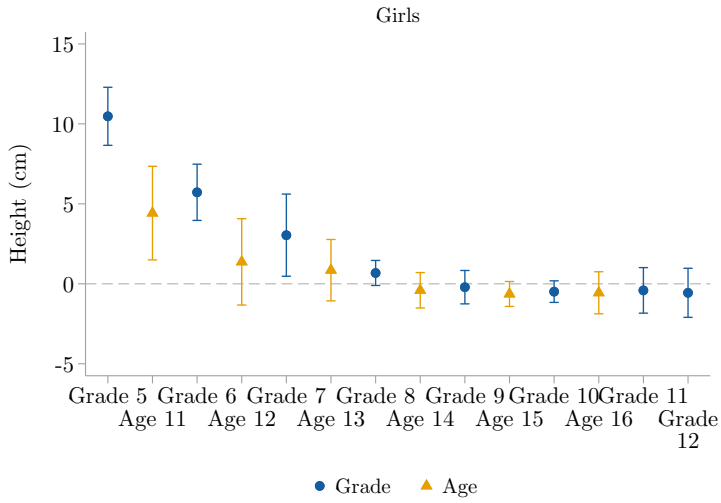
(a) Girls



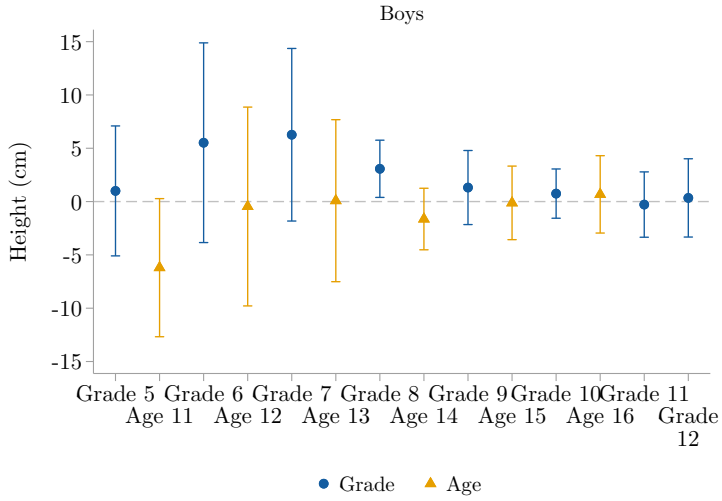
(b) Boys

The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent estimates using outcomes measured at the same grade. Triangles represent estimates using outcomes measured at the same age. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. Test scores are vertically scaled and standardized with mean zero and unit variance.

FIGURE 1.12. Old-for-grade Effects on Test Scores in Grades 7–12 and Ages 13–17



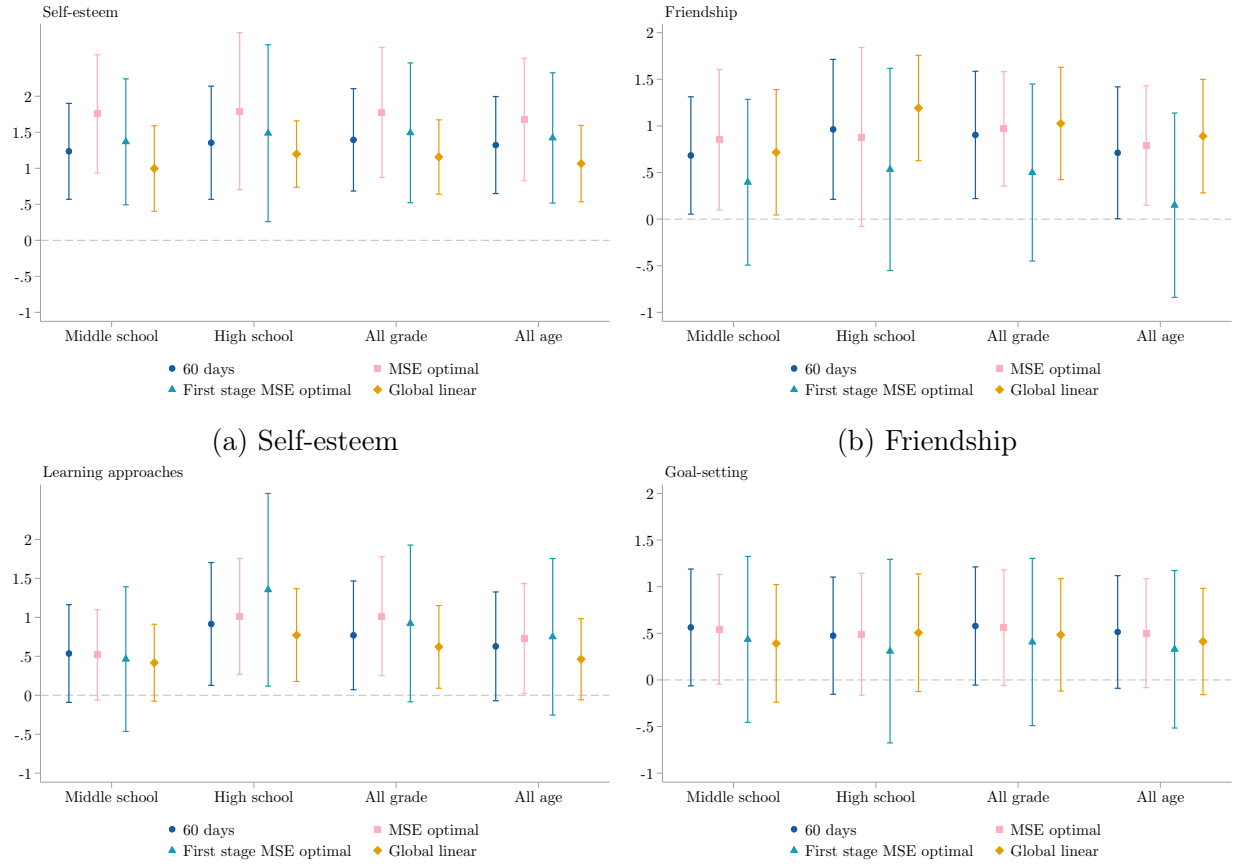
(a) Girls



(b) Boys

The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent estimates using outcomes measured at the same grade. Triangles represent estimates using outcomes measured at the same age. The sample includes January- to April-born children among fourth-grade and seventh-grade cohorts in KCYPS 2010. All regressions include cohort fixed effects. Standard errors are clustered at the month of birth level.

FIGURE 1.13. Old-for-grade Effects on Height in Grades 5–12 and Ages 11–16 Using IV



(a) Self-esteem

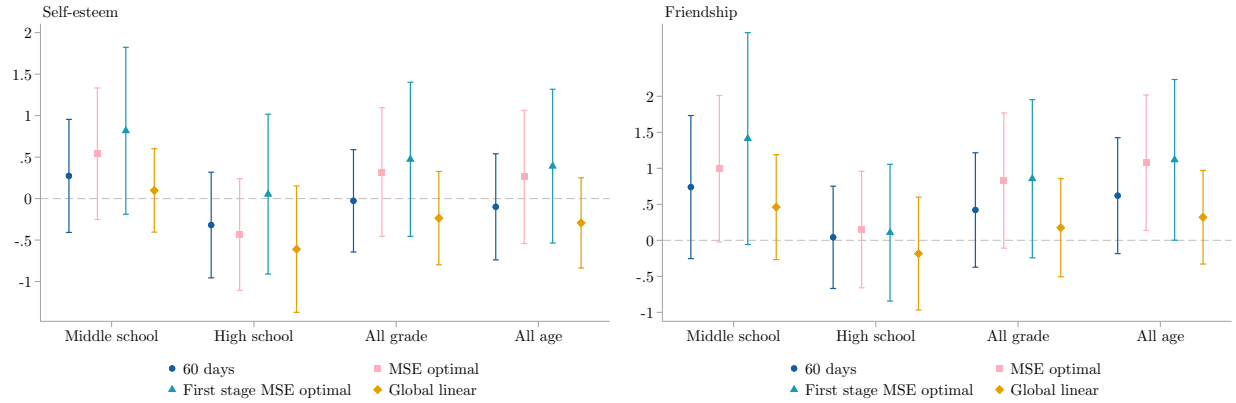
(b) Friendship

(c) Learning approaches

(d) Goal-setting mindsets

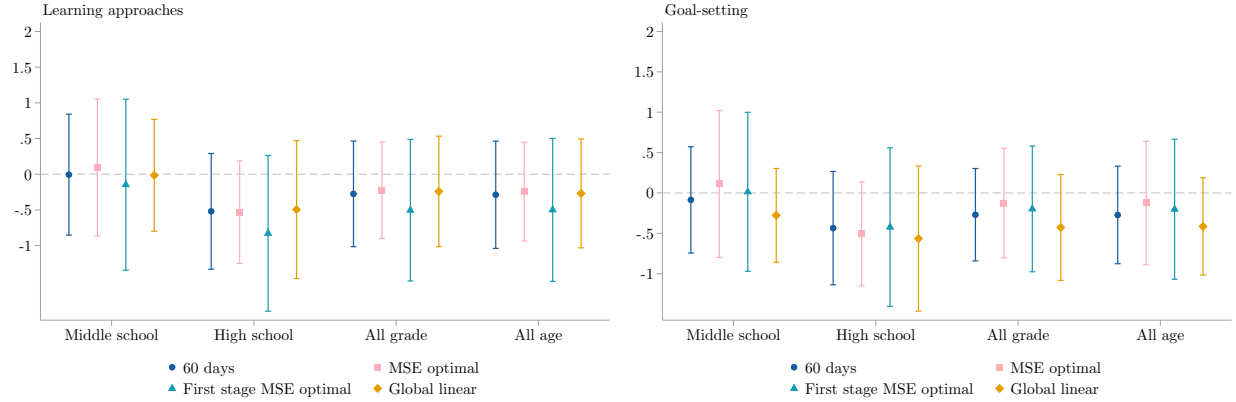
The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent estimates from the main regression using 60-day bandwidths. Squares represent bias-corrected estimates and robust confidence intervals from the regression with mean squared error (MSE) optimal bandwidths following Calonico et al. (2014). Triangles represent bias-corrected estimates and robust confidence intervals from the regression with first-stage MSE optimal bandwidths. Lastly, diamonds represent estimates using global linear polynomials with 60-day bandwidths. All estimates use a triangular kernel and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.14. Sensitivity of RD Estimates to Bandwidth Choices: Being Old for Grade on Social-Emotional Skills (Girls)



(a) Self-esteem

(b) Friendship

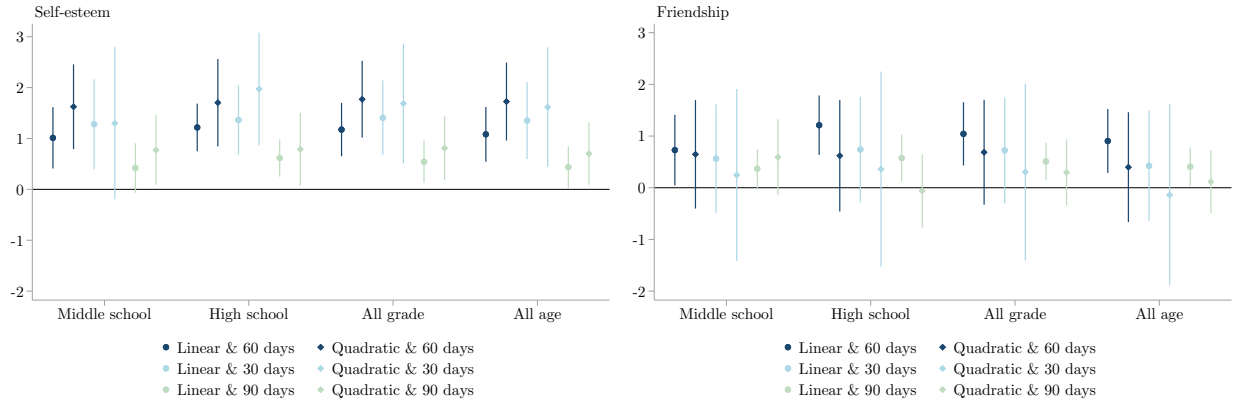


(c) Learning appraches

(d) Goal-setting mindsets

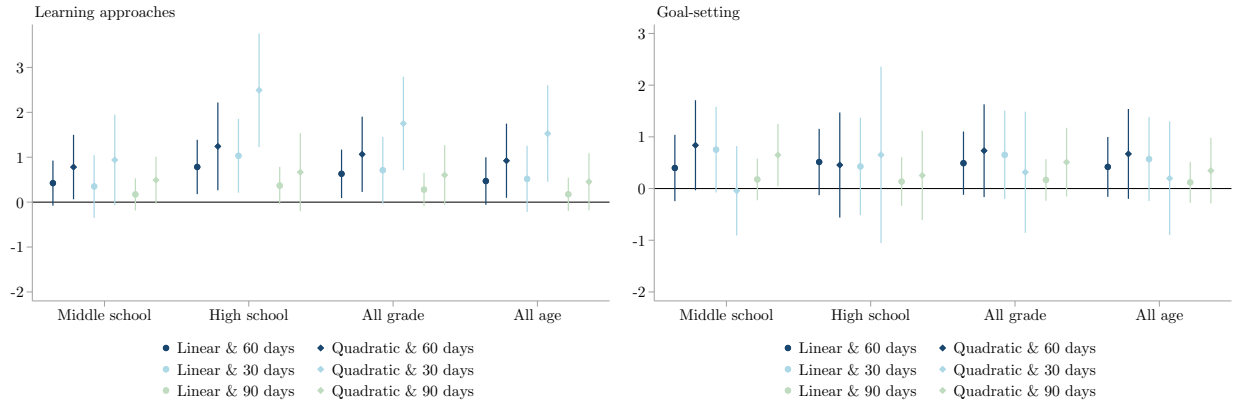
The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent estimates from the main regression using 60-day bandwidths. Squares represent bias-corrected estimates and robust confidence intervals from the regression with mean squared error (MSE) optimal bandwidths following Calonico et al. (2014). Triangles represent bias-corrected estimates and robust confidence intervals from the regression with first-stage MSE optimal bandwidths. Lastly, diamonds represent estimates using global linear polynomials with 60-day bandwidths. All estimates use a triangular kernel and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.15. Sensitivity of RD Estimates to Bandwidth Choices: Being Old for Grade on Social-Emotional Skills (Boys)



(a) Self-esteem

(b) friendship

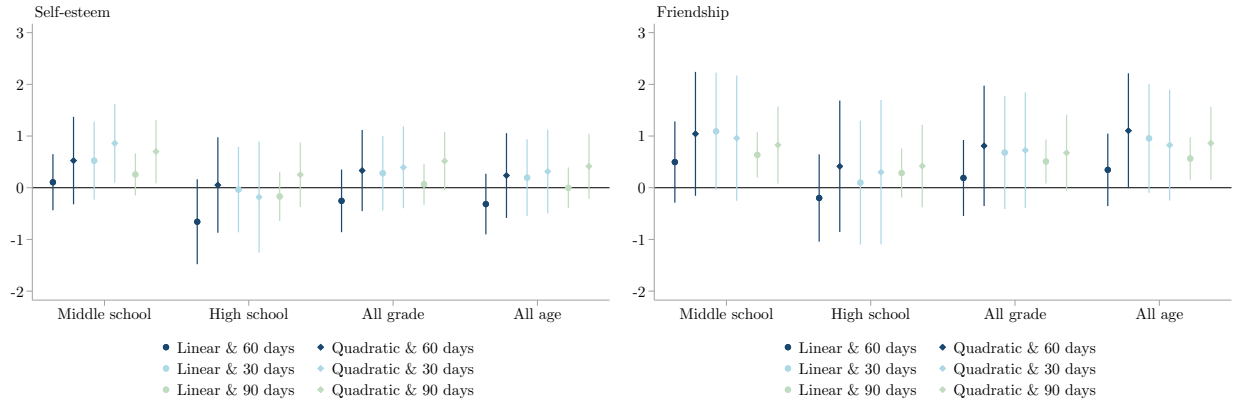


(c) Learning

(d) Goal-setting mindsets

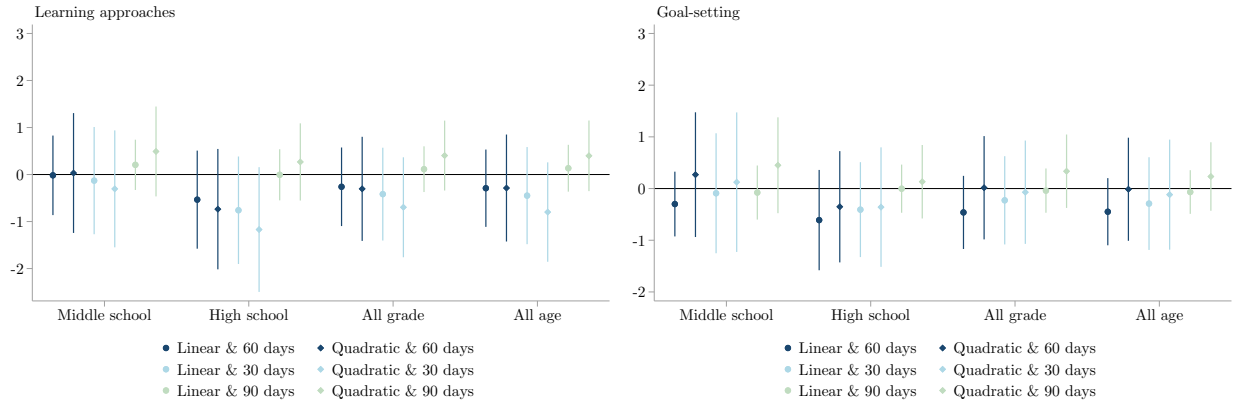
The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent the estimates using global linear polynomials. Diamonds represent the estimates using global quadratic polynomials. All estimates include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.16. Sensitivity of RD Estimates to Polynomial and Bandwidth Choices: Being Old for Grade on Social-Emotional Skills (Girls)



(a) Self-esteem

(b) friendship

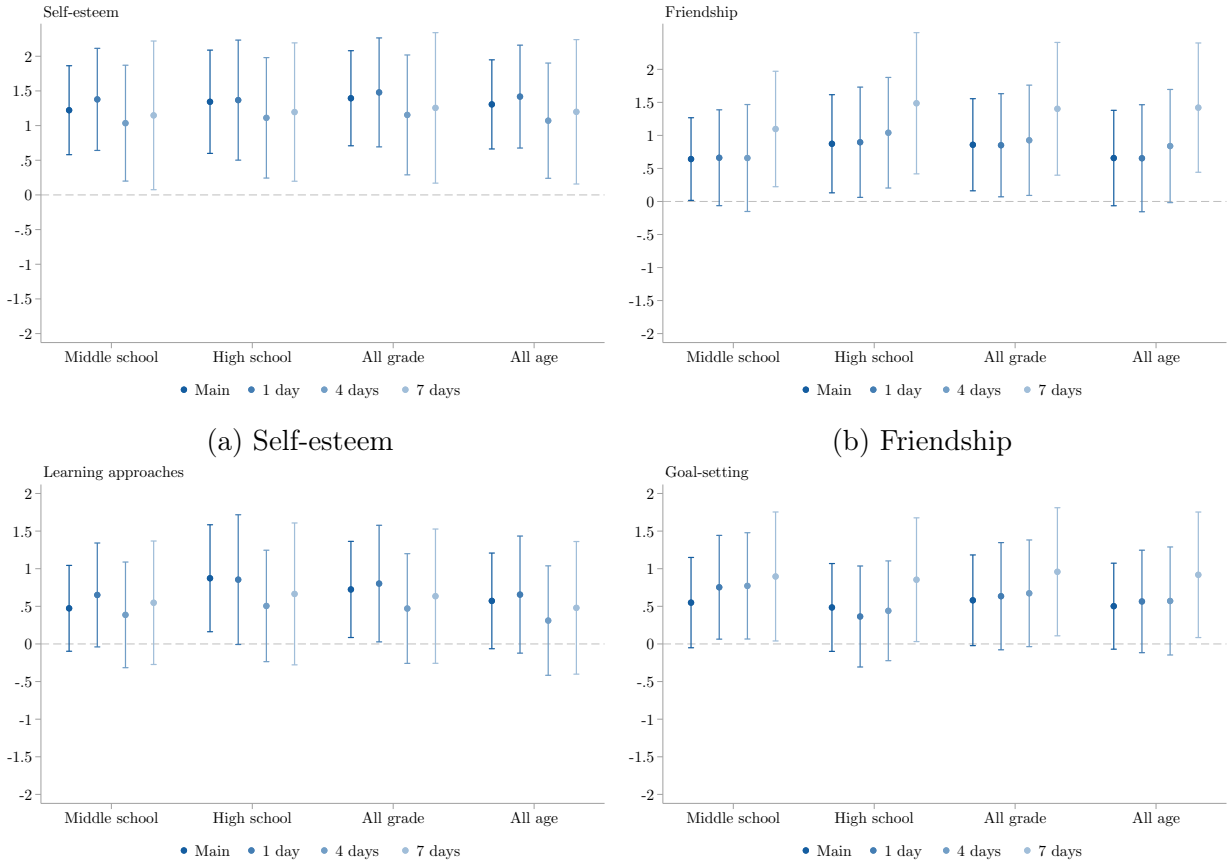


(c) Learning

(d) Goal-setting mindsets

The estimates and corresponding 95 percent confidence intervals (CI) are shown. Circles represent the estimates using global linear polynomials. Diamonds represent the estimates using global quadratic polynomials. All estimates include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.17. Sensitivity of RD Estimates to Polynomial and Bandwidth Choices: Being Old for Grade on Social-Emotional Skills (Boys)



(a) Self-esteem

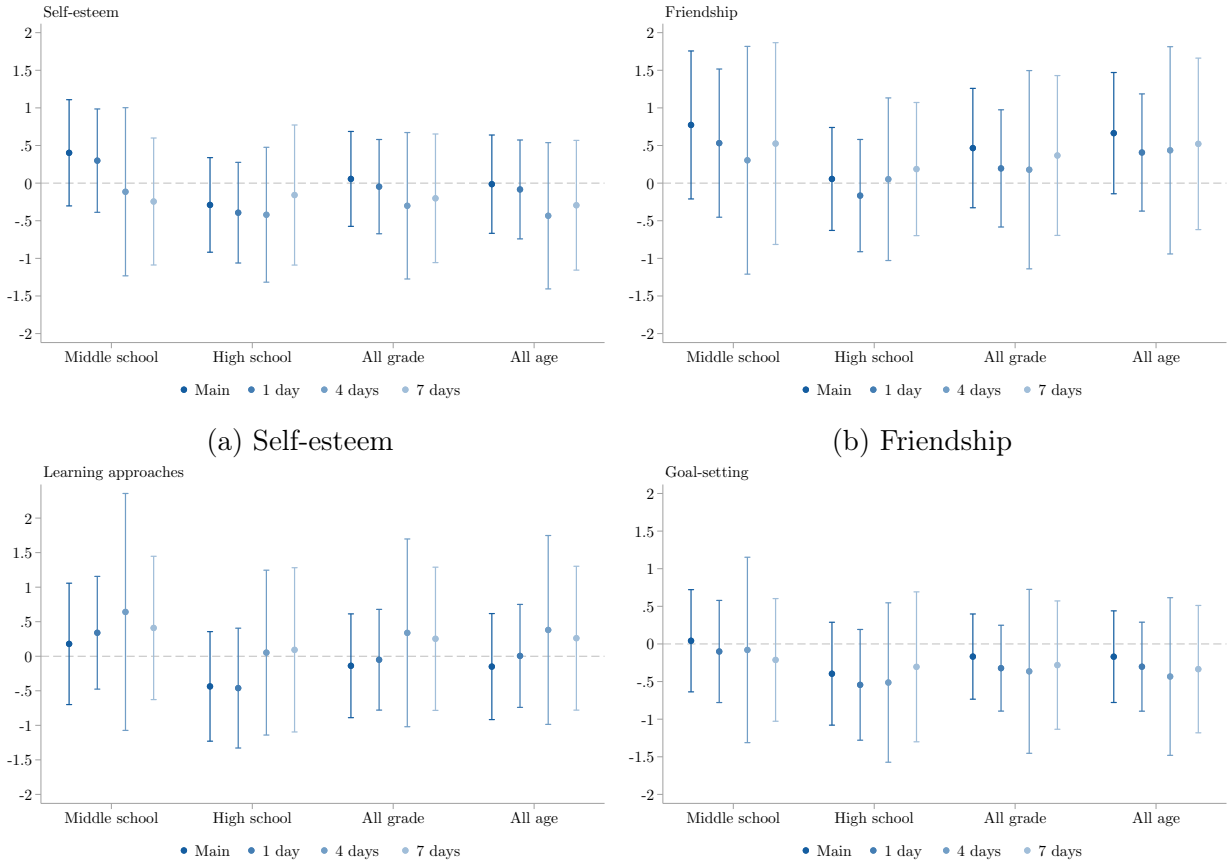
(b) Friendship

(c) Learning approaches

(d) Goal-setting mindsets

The estimates and corresponding 95 percent confidence intervals (CI) are shown. Donut sizes are set to 1 day, 4 days, and 7 days before and after the cutoff. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.18. Comparison of Standard RD and Donut RD Estimates: Effects of Being Old for Grade on Social-Emotional Skills (Girls)



(a) Self-esteem

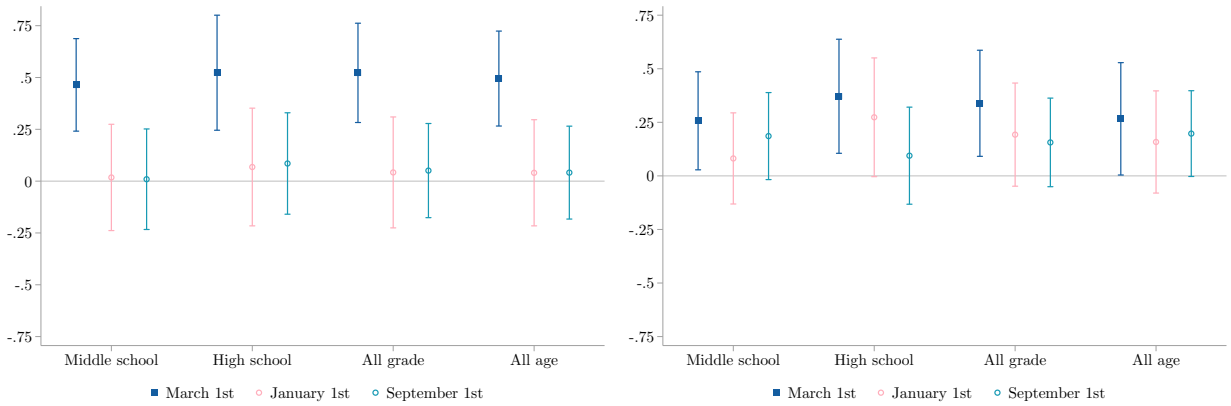
(b) Friendship

(c) Learning approaches

(d) Goal-setting mindsets

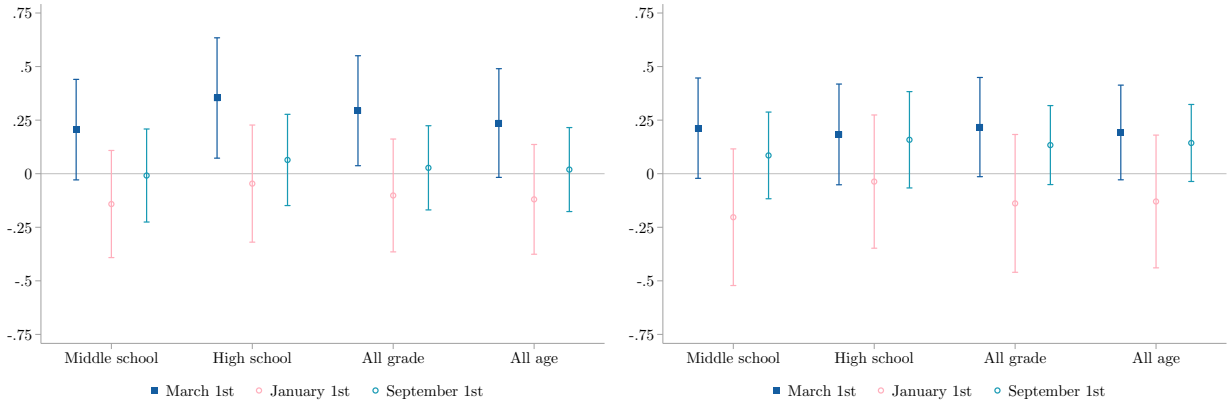
The estimates and corresponding 95 percent confidence intervals (CI) are shown. Donut sizes are set to 1 day, 4 days, and 7 days before and after the cutoff. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.19. Comparison of Standard RD and Donut RD Estimates: Effects of Being Old for Grade on Social-Emotional Skills (Boys)



(a) Self-esteem

(b) Friendship

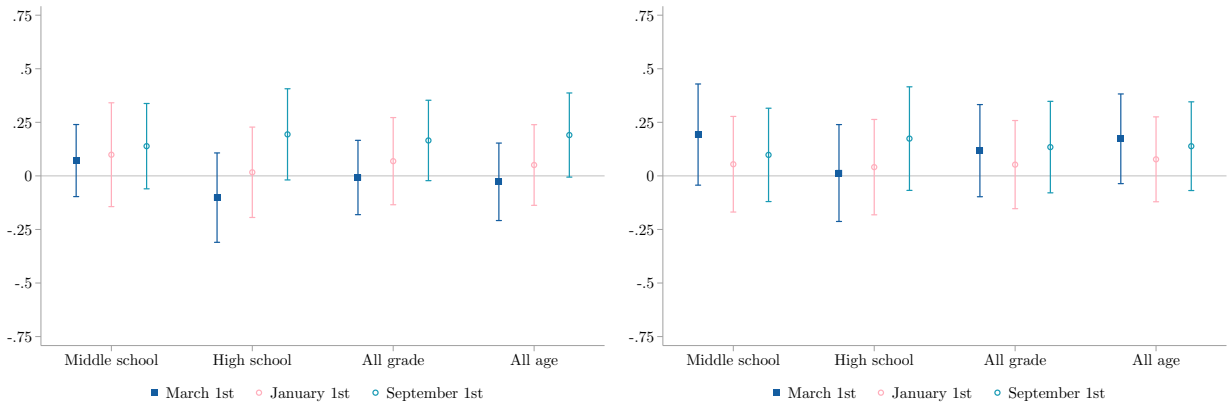


(c) Learning approaches

(d) Goal-setting mindsets

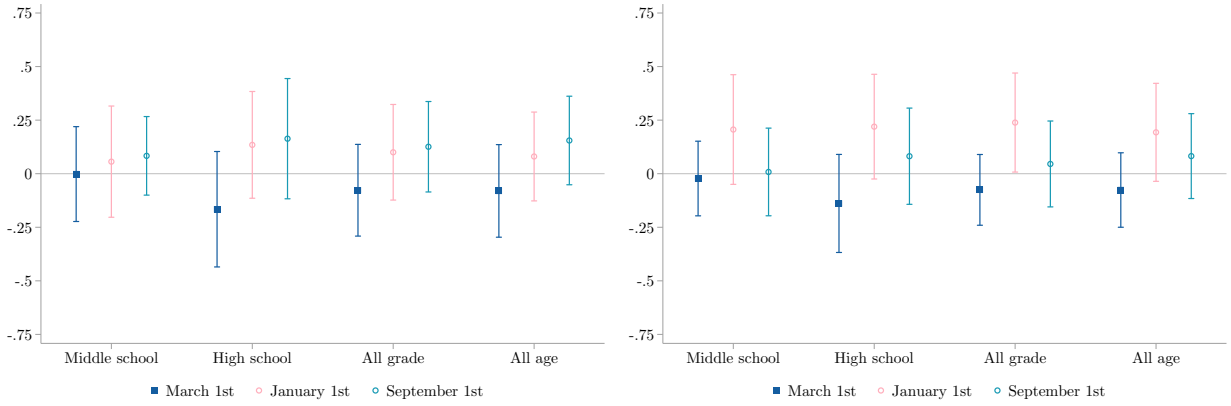
The reduced-form estimates and corresponding 95 percent confidence intervals (CI) are shown. Squares represent estimates from the main specification. Hollow circles represent estimates using January 1st and September 1st as a placebo cutoff. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.20. Comparison of Reduced-Form RD Estimates with Placebo Age Cutoffs: Effects of Being Born after the Cutoff on Social-Emotional Skills (Girls)



(a) Self-esteem

(b) Friendship



(c) Learning appraches

(d) Goal-setting mindsets

The reduced-form estimates and corresponding 95 percent confidence intervals (CI) are shown. Squares represent estimates from the main specification. Hollow circles represent estimates using January 1st and September 1st as a placebo cutoff. All regressions use a triangular kernel, linear polynomials, and 60-day bandwidths and include cohort fixed effects. Standard errors are clustered at the date of birth level. All outcomes are standardized with mean zero and unit variance.

FIGURE 1.21. Comparison of Reduced-Form RD Estimates with Placebo Age Cutoffs: Effects of Being Born after the Cutoff on Social-Emotional Skills (Boys)

1.8. Tables

TABLE 1.1. Summary Statistics of Outcome Indices by Cutoff and Gender: January to April Borns

	Girls		Boys	
	Before cutoff	After cutoff	Before cutoff	After cutoff
Panel A: Self-esteem				
Grades 7 to 9	3.597	3.699	3.728	3.718
	(0.707)	(0.735)	(0.683)	(0.666)
Grades 10 to 12	3.623	3.667	3.807	3.778
	(0.705)	(0.694)	(0.665)	(0.701)
Panel B: Friendship				
Grades 7 to 9	4.221	4.264	4.206	4.265
	(0.497)	(0.497)	(0.557)	(0.532)
Grades 10 to 12	4.185	4.209	4.231	4.212
	(0.508)	(0.528)	(0.525)	(0.536)
Panel C: Learning approaches				
Grades 7 to 9	3.275	3.342	3.333	3.339
	(0.635)	(0.622)	(0.650)	(0.622)
Grades 10 to 12	3.314	3.351	3.336	3.315
	(0.633)	(0.605)	(0.605)	(0.634)
Panel D: Goal-setting				
Grades 7 to 9	3.695	3.744	3.791	3.745
	(0.676)	(0.698)	(0.662)	(0.666)
Grades 10 to 12	3.704	3.748	3.731	3.693
	(0.625)	(0.630)	(0.631)	(0.701)
Observations	648	648	700	809

Average and standard deviation before standardization to mean zero and a unit variance are shown, with standard deviations in parentheses. The sample I use includes fourth-grade and seventh-grade cohorts born between January and April from SELS 2010 data. Grades 7 to 9 and grades 10 to 12 comprise middle school and high school, respectively. For each survey question, the student response is on a 5-point scale: *1 'strongly disagree,' 2 'disagree,' 3 'neutral,' 4 'agree,' 5 'strongly agree'.* A higher index indicates better social-emotional skills.

TABLE 1.2. First Stage: Effects of Being Born after the Cutoff on Being Old for Grade

	Entire sample		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)
Born after cutoff	0.323***	0.315***	0.376***	0.372***	0.278***	0.277***
	(0.026)	(0.026)	(0.034)	(0.033)	(0.048)	(0.048)
Observations	2825	2825	1287	1287	1494	1494
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the estimate of being born after the cutoff on being old for grade from a separate regression with a triangular kernel, linear polynomials, and 60-day bandwidths. All regressions include cohort fixed effects. Standard errors, in parentheses, are clustered at the date of birth level. Individual-level controls include a female indicator, a firstborn indicator, parents' educational attainment (categorized as less than high school, high school graduates, and some college or more), and indicators for missing values of these controls.

TABLE 1.3. RD Estimates of Social-Emotional Skills in Middle School and High School

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) MS	(2) HS	(3) MS	(4) HS	(5) MS	(6) HS	(7) MS	(8) HS
Panel A: Girls								
Old for grade	1.237***	1.356***	0.683**	0.964**	0.536*	0.916**	0.563*	0.475
	(0.340)	(0.401)	(0.321)	(0.383)	(0.320)	(0.403)	(0.320)	(0.321)
Observations	1216	1082	1216	1082	1214	1082	1217	1082
Panel B: Boys								
Old for grade	0.275	-0.320	0.743	0.042	-0.007	-0.523	-0.087	-0.438
	(0.349)	(0.326)	(0.510)	(0.364)	(0.435)	(0.415)	(0.338)	(0.360)
Observations	1416	1206	1417	1206	1417	1206	1417	1206
Panel C: Entire Sample								
Old for grade	0.823***	0.590*	0.667**	0.485*	0.239	0.164	0.251	-0.027
	(0.271)	(0.317)	(0.318)	(0.280)	(0.279)	(0.304)	(0.179)	(0.262)
Observations	2675	2324	2676	2324	2674	2324	2677	2324
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the estimate of being old for grade on a social-emotional skill index from a separate regression with a triangular kernel, linear polynomials, and 60-day bandwidths. Standard errors, in parentheses, are clustered at the date of birth level. All regressions include cohort fixed effects. Middle school (MS) outcomes are averaged across grades 7 to 9, and high school (HS) outcomes are averaged across grades 10 to 12. All outcomes are standardized with mean zero and unit variance.

TABLE 1.4. Comparison of RD Estimates: Social-Emotional Skills Measured at Same Grades vs. at Same Ages

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) Grade	(2) Age	(3) Grade	(4) Age	(5) Grade	(6) Age	(7) Grade	(8) Age
Panel A: Girls								
Old for grade	1.396***	1.324***	0.904***	0.712**	0.770**	0.628*	0.579*	0.515*
	(0.363)	(0.344)	(0.349)	(0.361)	(0.356)	(0.357)	(0.324)	(0.308)
Observations	1286	1274	1286	1274	1284	1273	1287	1274
Panel B: Boys								
Old for grade	-0.027	-0.099	0.423	0.624	-0.277	-0.289	-0.271	-0.274
	(0.316)	(0.328)	(0.407)	(0.413)	(0.379)	(0.385)	(0.293)	(0.310)
Observations	1493	1483	1494	1484	1494	1484	1494	1484
Panel C: Entire Sample								
Old for grade	0.762***	0.685**	0.635**	0.635**	0.217	0.131	0.130	0.0998
	(0.283)	(0.276)	(0.297)	(0.298)	(0.274)	(0.276)	(0.192)	(0.187)
Observations	2823	2801	2824	2802	2822	2801	2825	2802
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the estimate of being old-for-grade on social-emotional skill indices from a separate regression with a triangular kernel, linear polynomials, and 60-day bandwidths. Standard errors, in parentheses, are clustered at the date of birth level. All regressions include cohort fixed effects. All grade outcome indices are averaged across grades 7 to 12, and all age outcome indices are averaged across ages 13 to 17. All outcomes are standardized with mean zero and unit variance.

TABLE 1.5. Average Outcome Indices by Cutoff and Gender: January to April Borns

	Girls		Boys	
	Before cutoff	After cutoff	Before cutoff	After cutoff
Panel A: Self-esteem				
Grades 7 to 12	3.611	3.686	3.749	3.744
	(0.658)	(0.664)	(0.637)	(0.636)
Ages 13 to 17	3.644	3.692	3.760	3.740
	(0.666)	(0.683)	(0.649)	(0.643)
Panel B: Friendship				
Grades 7 to 12	4.201	4.236	4.217	4.236
	(0.463)	(0.479)	(0.500)	(0.499)
Ages 13 to 17	4.211	4.244	4.226	4.251
	(0.476)	(0.479)	(0.505)	(0.499)
Panel C: Learning approaches				
Grades 7 to 12	3.296	3.355	3.329	3.325
	(0.606)	(0.575)	(0.591)	(0.587)
Ages 13 to 17	3.297	3.351	3.322	3.319
	(0.615)	(0.591)	(0.615)	(0.598)
Panel D: Goal-setting				
Grades 7 to 12	3.714	3.753	3.759	3.729
	(0.603)	(0.632)	(0.587)	(0.630)
Ages 13 to 17	3.706	3.743	3.739	3.723
	(0.617)	(0.647)	(0.616)	(0.636)
Observations	648	648	700	809

Average and standard deviation before standardization to mean zero and a unit variance are shown, with standard deviations in parentheses. The sample I use includes fourth-grade and seventh-grade cohorts born between January and April from SELS 2010 data. For each survey question, the student response is on a 5-point scale: *1 'strongly disagree,' 2 'disagree,' 3 'neutral,' 4 'agree,' 5 'strongly agree'*. A higher index indicates better social-emotional skills.

TABLE 1.6. Root Mean Squared Error for Each Outcome in Middle School and High School

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) MS	(2) HS	(3) MS	(4) HS	(5) MS	(6) HS	(7) MS	(8) HS
Panel A: Girls								
Linear	0.155	0.183	0.146	0.288	0.115	0.331	0.127	0.128
Quadratic	0.379	0.467	0.250	0.317	0.359	1.196	0.490	0.430
Cubic	1.095	1.229	1.170	0.934	1.734	2.299	1.017	1.228
Panel B: Boys								
Linear	0.425	0.277	0.583	0.168	0.215	0.217	0.229	0.168
Quadratic	0.503	0.405	0.510	0.321	0.783	0.961	0.443	0.441
Cubic	0.803	0.929	0.854	1.201	1.094	1.223	0.704	1.225

Each cell reports a result from a separate regression with a triangular kernel and 60-day bandwidths. Middle school (MS) outcomes are averaged across grades 7 to 9, high school (HS) outcomes are averaged across grades 10 to 12.

TABLE 1.7. Root Mean Squared Error for Each Outcome in All Grades and All Ages

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) Grade	(2) Age	(3) Grade	(4) Age	(5) Grade	(6) Age	(7) Grade	(8) Age
Panel A: Girls								
Linear	0.159	0.158	0.236	0.359	0.158	0.147	0.129	0.127
Quadratic	0.371	0.350	0.249	0.262	0.636	0.547	0.317	0.312
Cubic	1.085	0.993	0.981	0.870	1.989	1.886	0.871	0.784
Panel B: Boys								
Linear	0.353	0.354	0.349	0.411	0.188	0.191	0.171	0.171
Quadratic	0.370	0.356	0.394	0.406	0.880	1.011	0.406	0.378
Cubic	0.827	0.796	0.852	0.848	1.198	1.355	0.895	0.841

Each cell reports a result from a separate regression with a triangular kernel and 60-day bandwidths. All grade outcomes are averaged across grades 7 to 12. All age outcomes are averaged across ages 13 to 17.

TABLE 1.8. First Stage using MSE Optimal Bandwidths: Effects of Being Born after the Cutoff on Being Old for Grade

	Entire sample			Girls			Boys		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Born after cutoff	0.295	0.283	0.294	0.359	0.359	0.357	0.271	0.270	0.271
	[.23 ; .359]	[.217 ; .35]	[.229 ; .358]	[.274 ; .444]	[.28 ; .439]	[.271 ; .443]	[.127 ; .414]	[.129 ; .411]	[.129 ; .413]
Bandwidth	31	32	31	38	38	37	40	40	40
Observations	1,505	1,548	1,505	834	834	813	1,014	1,014	1,014
Cohort FE	Yes	Yes		Yes	Yes		Yes	Yes	
Individual controls		Yes			Yes			Yes	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents bias-corrected estimates of being born after the cutoff on being old for grade and robust 95 percent confidence intervals using the mean squared error optimal bandwidths. Each cell reports results from a separate regression with a triangular kernel and linear polynomials. Standard errors are clustered at the date of birth level. Individual-level controls include a female indicator, a firstborn indicator, parents' educational attainment (categorized as less than high school, high school graduates, and some college or more), and indicators for missing values of these controls.

TABLE 1.9. RD Estimates using MSE Optimal Bandwidths: Social-Emotional Skills in Middle School and High School

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) MS	(2) HS	(3) MS	(4) HS	(5) MS	(6) HS	(7) MS	(8) HS
Panel A: Girls								
Old for grade	1.721	1.785	0.817	1.003	0.545	1.055	0.571	0.511
	[.902 ; 2.54]	[.704 ; 2.865]	[.049 ; 1.584]	[.068 ; 1.938]	[-.054 ; 1.143]	[.293 ; 1.817]	[-.035 ; 1.177]	[-.159 ; 1.182]
Bandwidth	51	48	54	49	73	70	72	66
Observations	1,035	870	1,113	886	1,487	1,273	1,478	1,203
Panel B: Boys								
Old for grade	0.517	-0.492	1.055	0.147	0.097	-0.571	0.114	-0.547
	[-.264 ; 1.297]	[-1.167 ; .184]	[-.008 ; 2.119]	[-.666 ; .96]	[-.886 ; 1.08]	[-1.317 ; .175]	[-.793 ; 1.021]	[-1.215 ; .122]
Bandwidth	56	69	66	60	62	78	42	77
Observations	1,322	1,396	1,549	1,224	1,471	1,546	1,000	1,538
Panel C: Entire Sample								
Old for grade	1.246	0.904	0.981	0.537	0.369	0.199	0.343	-0.078
	[.573 ; 1.919]	[.083 ; 1.725]	[.207 ; 1.756]	[.004 ; 1.07]	[-.283 ; 1.022]	[-.358 ; .757]	[-.088 ; .774]	[-.575 ; .419]
Bandwidth	42	46	55	72	61	74	42	73
Observations	1,891	1,800	2,475	2,789	2,752	2,862	1,893	2,829
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the estimate of being old for grade on a social-emotional skill index and robust 95 percent confidence intervals from a separate regression with a triangular kernel, linear polynomials, and mean squared error optimal bandwidths. Standard errors are clustered at the date of birth level. All regressions include cohort fixed effects. Middle school (MS) outcomes are averaged across grades 7 to 9, high school (HS) outcomes are averaged across grades 10 to 12. All outcomes are standardized with mean zero and unit variance.

TABLE 1.10. Comparison of RD Estimates using MSE Optimal Bandwidths: Social-Emotional Skills Measured in All Grades and All Ages

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) Grade	(2) Age	(3) Grade	(4) Age	(5) Grade	(6) Age	(7) Grade	(8) Age
Panel A: Girls								
Old for grade	1.742	1.643	1.013	0.829	1.083	0.860	0.589	0.528
	[.84 ; 2.644]	[.792 ; 2.493]	[.382 ; 1.645]	[.168 ; 1.491]	[.29 ; 1.875]	[.078 ; 1.642]	[-.05 ; 1.228]	[-.075 ; 1.132]
Bandwidth	47	47	75	75	62	63	70	70
Observations	1,016	1,007	1,628	1,616	1,353	1,354	1,521	1,508
Panel B: Boys								
Old for grade	0.247	0.214	0.813	1.063	-0.243	-0.264	-0.123	-0.120
	[-.498 ; .992]	[-.567 ; .996]	[-.123 ; 1.749]	[.127 ; 2]	[-.945 ; .458]	[-.983 ; .455]	[-.801 ; .554]	[-.885 ; .645]
Bandwidth	50	47	52	53	75	75	46	44
Observations	1,246	1,178	1,296	1,306	1,825	1,814	1,161	1,102
Panel C: Entire Sample								
Old for grade	1.132	1.042	0.806	0.811	0.301	0.220	0.235	0.189
	[.4 ; 1.864]	[.351 ; 1.734]	[.202 ; 1.411]	[.201 ; 1.421]	[-.233 ; .835]	[-.357 ; .797]	[-.205 ; .674]	[-.252 ; .629]
Bandwidth	41	41	68	68	71	68	51	49
Observations	1,965	1,952	3,210	3,188	3,324	3,187	2,407	2,304
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell reports the estimate of being old for grade on a social-emotional skill index and robust 95 percent confidence intervals from a separate regression with a triangular kernel, linear polynomials, and mean squared error optimal bandwidths. Standard errors are clustered at the date of birth level. All regressions include cohort fixed effects. All grade outcome indices are averaged across grades 7 to 12, and all age outcome indices are averaged across ages 13 to 17. All outcomes are standardized with mean zero and unit variance.

TABLE 1.11. Reduced-Form: Effects of Being Born after the Cutoff on Social-Emotional Skills in Middle School and High School

	Self-esteem		Friendship		Learning		Goal-setting	
	(1) MS	(2) HS	(3) MS	(4) HS	(5) MS	(6) HS	(7) MS	(8) HS
Panel A: Girls								
Born after cutoff	0.464***	0.523***	0.257**	0.372***	0.206*	0.353**	0.213*	0.183
	(0.114)	(0.142)	(0.117)	(0.136)	(0.120)	(0.143)	(0.119)	(0.120)
Observations	1,216	1,082	1,216	1,082	1,214	1,082	1,217	1,082
Panel B: Boys								
Born after cutoff	0.072	-0.101	0.193	0.013	-0.002	-0.166	-0.022	-0.139
	(0.086)	(0.106)	(0.120)	(0.115)	(0.113)	(0.137)	(0.089)	(0.117)
Observations	1,416	1,206	1,417	1,206	1,417	1,206	1,417	1,206
Panel C: Entire Sample								
Born after cutoff	0.261***	0.201*	0.214**	0.169*	0.079	0.055	0.083	-0.011
	(0.077)	(0.105)	(0.093)	(0.095)	(0.086)	(0.103)	(0.055)	(0.090)
Observations	2,675	2,324	2,676	2,324	2,674	2,324	2,677	2,324
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school (MS) outcomes are averaged across grades 7 to 9, high school (HS) outcomes are averaged across grades 10 to 12. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.12. Effects of Being Old for Grade on Test Scores in Middle and High School

	Korean		English		Math		
	(1) MS	(2) HS	(3) MS	(4) HS	(5) MS	(6) General	(7) Science
Panel A: Girls							
Old for grade	0.260	0.105	-0.099	-0.091	0.286	0.001	-0.500
	(0.261)	(0.321)	(0.313)	(0.398)	(0.305)	(0.359)	(0.611)
Observations	1,227	1,081	1,227	1,081	1,227	843	385
Panel B: Boys							
Old for grade	0.163	0.038	0.593	0.471*	0.125	-0.406	0.028
	(0.409)	(0.299)	(0.369)	(0.272)	(0.391)	(0.447)	(0.534)
Observations	1,441	1,203	1,441	1,202	1,441	738	678
Panel C: Entire Sample							
Old for grade	0.151	0.031	0.192	0.150	0.230	-0.127	-0.140
	(0.263)	(0.210)	(0.258)	(0.235)	(0.261)	(0.278)	(0.374)
Observations	2,712	2,320	2,712	2,319	2,712	1,610	1,081
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.13. Effects of Being Old for Grade on Height, Weight, and BMI

	Fourth-grade cohort			Seventh-grade cohort		
	(1) Height(cm)	(2) Weight(kg)	(3) BMI	(4) Height(cm)	(5) Weight(kg)	(6) BMI
Panel A: Girls						
Old for grade	7.047	-4.762	-4.041	1.237	3.231	1.013
	(4.502)	(6.878)	(2.582)	(1.623)	(2.709)	(0.851)
Baseline mean	137.335	33.913	17.871	156.410	46.057	18.786
Observations	584	584	584	624	580	580
Panel B: Boys						
Old for grade	4.588	3.067	0.478	0.325	2.005	0.455
	(3.346)	(4.099)	(1.428)	(5.019)	(10.253)	(3.226)
Baseline mean	137.119	35.499	18.781	156.805	48.921	19.769
Observations	621	621	621	788	768	768

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. 1cm = 0.3937in & 1kg = 2.205lb.

TABLE 1.14. RD Estimates Robust to Inclusion of Cohort Fixed Effects, Controls, and Day-of-Week Fixed Effects: Self-Esteem

	Middle school				High school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
Old for grade	1.236***	1.180***	1.229***	1.276***	1.355***	1.306***	1.311***	1.411***
	(0.340)	(0.327)	(0.334)	(0.338)	(0.401)	(0.382)	(0.386)	(0.386)
Observations	1216	1216	1216	1216	1082	1082	1082	1082
Panel B: Boys								
Old for grade	0.273	0.401	0.271	0.235	-0.319	-0.250	-0.315	-0.296
	(0.348)	(0.355)	(0.344)	(0.314)	(0.325)	(0.328)	(0.321)	(0.309)
Observations	1416	1416	1416	1416	1206	1206	1206	1206
Cohort FE	Yes	Yes		Yes	Yes	Yes		Yes
Individual controls		Yes				Yes		
Day-of-week FE				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school outcomes are averaged across 7th-9th grades, and high school outcomes are averaged across 10th-12th grades. All outcomes are standardized with mean zero and unit variance.

TABLE 1.15. RD Estimates Robust to Inclusion of Cohort Fixed Effects, Controls, and Day-of-Week Fixed Effects: Friendship

	Middle school				High school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
Old for grade	0.683**	0.624*	0.686**	0.718**	0.963**	0.858**	0.967**	1.017***
	(0.321)	(0.319)	(0.316)	(0.323)	(0.383)	(0.380)	(0.377)	(0.386)
Observations	1216	1216	1216	1216	1082	1082	1082	1082
Panel B: Boys								
Old for grade	0.739	0.768	0.746	0.702	0.042	0.068	0.031	0.035
	(0.507)	(0.495)	(0.504)	(0.499)	(0.362)	(0.345)	(0.357)	(0.352)
Observations	1417	1417	1417	1417	1206	1206	1206	1206
Cohort FE	Yes	Yes		Yes	Yes	Yes		Yes
Individual controls		Yes				Yes		
Day-of-week FE				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school outcomes are averaged across 7th-9th grades, and high school outcomes are averaged across 10th-12th grades. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.16. RD Estimates Robust to Inclusion of Cohort Fixed Effects, Controls, and Day-of-Week Fixed Effects: Learning Approaches

	Middle school				High school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
Old for grade	0.536*	0.431	0.551*	0.557*	0.915**	0.829**	0.899**	0.964**
	(0.320)	(0.284)	(0.316)	(0.312)	(0.402)	(0.356)	(0.393)	(0.396)
Observations	1214	1214	1214	1214	1082	1082	1082	1082
Panel B: Boys								
Old for grade	-0.007	0.181	-0.008	-0.033	-0.521	-0.407	-0.524	-0.504
	(0.432)	(0.436)	(0.428)	(0.419)	(0.414)	(0.397)	(0.409)	(0.377)
Observations	1417	1417	1417	1417	1206	1206	1206	1206
Cohort FE	Yes	Yes		Yes	Yes	Yes		Yes
Individual controls		Yes				Yes		
Day-of-week FE				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school outcomes are averaged across 7th-9th grades, and high school outcomes are averaged across 10th-12th grades. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.17. RD Estimates Robust to Inclusion of Cohort Fixed Effects, Controls, and Day-of-Week Fixed Effects: Goal-Setting Mindsets

	Middle school				High school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Girls								
Old for grade	0.563*	0.510*	0.562*	0.581*	0.475	0.459	0.438	0.504
	(0.320)	(0.309)	(0.315)	(0.322)	(0.321)	(0.299)	(0.305)	(0.321)
Observations	1217	1217	1217	1217	1082	1082	1082	1082
Panel B: Boys								
Old for grade	-0.086	0.047	-0.069	-0.082	-0.436	-0.383	-0.421	-0.397
	(0.336)	(0.346)	(0.335)	(0.331)	(0.358)	(0.346)	(0.352)	(0.310)
Observations	1417	1417	1417	1417	1206	1206	1206	1206
Cohort FE	Yes	Yes		Yes	Yes	Yes		Yes
Individual controls		Yes				Yes		
Day-of-week FE				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level. Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school outcomes are averaged across 7th-9th grades, and high school outcomes are averaged across 10th-12th grades. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.18. Robust to Inclusion of School Fixed Effects: Self-Esteem

	Middle school			High school			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Girls							
Old for grade	1.237***	1.237***	1.105***	1.356***	1.357***	0.971***	1.300***
	(0.340)	(0.336)	(0.219)	(0.401)	(0.388)	(0.223)	(0.263)
Observations	1,216	1,216	1,216	1,082	1,012	1,012	1,077
Panel B: Boys							
Old for grade	0.275	0.275	0.474	-0.320	-0.175	-0.109	-0.531*
	(0.349)	(0.422)	(0.340)	(0.326)	(0.434)	(0.319)	(0.295)
Observations	1,416	1,416	1,416	1,206	1,129	1,129	1,204
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Middle school FE			Yes			Yes	
High school FE							Yes
Cluster	DOB	MS	MS	DOB	MS	MS	HS

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered at the date of birth level (DOB), middle school level (MS), and high school level (HS). Each cell reports results from a separate regression with local linear polynomials, a triangular kernel, and a 60-day bandwidth. Middle school outcomes are averaged across 7th-9th grades, and high school outcomes are averaged across 10th-12th grades. All outcomes are standardized with mean zero and unit variance. All regressions include cohort fixed effects.

TABLE 1.19. Reliability Measures: Cronbach's α

	Middle school	High school	
		Academic	Vocational
Self-esteem	0.941	0.911	0.907
Friendship	0.782	0.785	0.787
Learning method	0.835	0.760	0.723
Learning effort	0.823	0.767	0.721
Learning attitude	0.776	0.717	0.655
Goal-setting	0.902	0.861	0.858

Cronbach's α is measured in the fourth survey wave.

CHAPTER 2

Do Single-Sex Middle Schools Improve Social-Emotional Skills

2.1. Introduction

Social-emotional skills, rarely captured by standardized tests aimed at achievement, play a crucial role in improving labor market outcomes. A substantial body of literature demonstrates the significant impact of these skills on employment, wages, and occupational choices.[1] Despite this, fewer studies focus on the factors that develop social-emotional skills compared to cognitive skills. Since social-emotional skills are primarily developed before adulthood, researchers have explored the role of schools (Lavy, 2010; McEachin et al., 2020; Jackson et al., 2020b), teachers (Gong et al., 2018; Kraft, 2019; Jackson, 2018), peers (Gong et al., 2021; Lavy and Sand, 2019), birth order (Black et al., 2018), and unconditional income transfers (Akee et al., 2018) in shaping these skills, but, there remains room to expand this line of research further.

This study examines the impact of single-sex schools on students' social-emotional skills, including self-esteem, learning approaches, goal-setting mindsets, social relationships, and school satisfaction, in addition to test scores. Single-sex schools offer a unique environment that can influence students through various channels, such as peer and teacher interactions. In coeducational schools, students can interact with the opposite gender through clubs or other activities, even if both genders are not in the same classroom. Previous studies have shown that having more female peers in the classroom improves the classroom environment and influences teacher behavior (Gong et al., 2021). At the extreme distribution of one

[1]Cunha and Heckman (2008); Cunha et al. (2010); Borghans et al. (2008a,c, 2014); Heckman et al. (2006); Heckman and Rubinstein (2001); Heckman and Kautz (2012); Lindqvist and Vestman (2011); Deming (2017); Kuhn and Weinberger (2005); Weinberger (2014); Bowles et al. (2001); Weidmann and Deming (2021); Edin et al. (2022); Sorrenti et al. (2025).

gender, these effects may be intensified. Moreover, teachers may behave differently with male and female students; for example, teacher behaviors may partly explain the differences in male and female students' test scores (Lavy, 2008). Understanding the impact of single-sex schools on social-emotional skills may help explain differences in later life outcomes.

I leverage random middle school assignments in Seoul, South Korea, to investigate the causal effect of single-sex schooling. Seoul implemented a random assignment policy in 1969 to reduce excessive competition in school entrance exams and alleviate educational inequality. Students are randomly assigned to one of the schools in the school sub-district of residence. This policy allows me to identify causal effects without selection bias which has complicated the study of single-sex schooling. Recent papers study the effects of single-sex schools on academic outcomes utilizing South Korean settings.[2]

I utilize the Seoul Education Longitudinal Study 2010 (SELS 2010), which provides comprehensive information on students' social-emotional skills compared to other datasets in South Korea. This enables an in-depth analysis of the impact of single-sex schools on social-emotional skills and test scores. Furthermore, this data allows me to control for elementary school effects and baseline outcome measures in sixth grade (the final year of elementary school). I include elementary-school fixed effects in the main specification to isolate skill development after entering middle school from skills built during elementary school. This is necessary because parents may sort their child into the elementary school, as students are typically assigned to their nearest school.

[2]There is evidence on single-sex secondary schools from other countries without random assignments, such as (Riordan, 1985; Lee and Bryk, 1986; Mael et al., 2005; Jackson, 2012; Pahlke et al., 2014). Mael et al. (2005); Pahlke et al. (2014) also cover elementary school levels. One paper evaluates an exogenous policy change in Trinidad and Tobago that converted 20 low-performing secondary schools from coeducational to single-sex in 2010 (Jackson, 2019). Boys show improved test scores and decreased arrest by ages 17 to 18 with single-sex schooling. There is a significant decrease in teen pregnancy by age 18 for girls.

Results show that girls and boys experience different effects from attending single-sex schools. Boys who attend all-boy middle schools for three years–compared to those in coeducational schools, exhibit higher self-esteem– better learning approaches, and higher Mathematics and English test scores. These improved learning approaches suggest that boys are more motivated and effective in their studies. Conversely, attending all-girl schools does not necessarily enhance girls' social-emotional skills. However, there is suggestive evidence of higher test scores for girls in all-girl middle schools, although these estimates are imprecise.

This study contributes to the literature on single-sex schooling by providing clean causal evidence about the effects of single-sex schools on social-emotional skills, utilizing random middle school assignments and unique panel survey data. Few studies have specifically focused on the impact of single-sex schools on social-emotional skills.[3] Some research has examined non-test score outcomes as mechanisms through which single-sex schooling affects academic performance in the Korean context.[4] One study investigates the impact on self-esteem and school aspirations of students attending single-sex school using a difference-in-differences approach with national data (Kim and Kim, 2022). In my research, leveraging random middle school assignments alleviates concerns about sorting into specific types of schools. For example, if students sort into schools, unobserved student characteristics may confound the estimated school effects. Additionally, I control for potential neighborhood sorting into elementary schools and elementary school effects on skill development by including elementary school fixed effects. Lastly, incorporating sixth-grade skill measures helps control for individual baseline achievement and isolate the impact of middle school.

I also contribute to the literature on the effects of single-sex schooling on cognitive skills. Previous research on single-sex school assignment in Seoul has shown mixed results regarding academic achievement. For instance, several studies have demonstrated that attending

[3]Lee and Nakazawa (2021) study the effect of single-sex schooling on labor market outcomes in South Korea. They find no significant effect for men, but single-sex schooling negatively affects women's labor market outcomes.

[4]For example, self-efficacy and interest in science as factors in choosing STEM majors (Park et al., 2018), and time-use and effort as factors affecting test scores (Lee et al., 2014).

single-sex high schools is significantly associated with higher scores on the College Scholastic Ability Test (CSAT) (Park et al., 2013; Choi et al., 2014; Dustmann et al., 2018).[5] Additionally, attending all-boy high schools is correlated with an increased likelihood of choosing a STEM major (Park et al., 2018), generally indicating positive effects of single-sex schooling on boys' academic outcomes. In contrast, Sohn (2016) shows no significant effects after accounting for neighborhood sorting and focusing on public school students. My findings show that single-sex schools improve students' test scores even after controlling for elementary school fixed effects and sixth-grade test scores.

The remainder of this paper is organized as follows. Section 2 describes education in South Korea and the background of the Equalization Policy. Section 3 explains the data and balance test, and Section 4 explains the estimation strategy. Then I show the results in section 5. Finally, section 6 provides the conclusion.

2.2. Education in South Korea and Equalization Policy in Seoul

Students in South Korea attend six years of elementary school, three years of middle school, and three years of high school. Elementary and middle school education is compulsory and free for all students. Most elementary and middle schools follow the national curriculum using government-issued or government-approved textbooks.[6] They are financed by the government and share the same qualification requirements and salary schedules for teachers. Students are assigned to the nearest public elementary school when they first enter.[7]

[5]The CSAT is a high-stakes exam required for university admission in South Korea, with approximately 70 percent of high school graduates pursuing higher education.

[6]Around 10 percent of schools have autonomous curricula, which allow more flexible curricula than other schools. Most of these schools are coeducational, so including these schools may attenuate single-sex school impacts. In my sample, 5 percent of students attend those schools. Excluding these schools also does not change the results.

[7]National and private elementary schools have a slightly different system. Parents first apply for one of the national or private elementary schools, then students are randomly selected afterward. There were two national elementary schools and 40 private elementary schools in Seoul in 2005. If students are not selected for national or private schools, they are also assigned to the nearest public elementary school.

Before the Equalization Policy was first enforced in Seoul in 1969, students were required to take an entrance exam to get into their desired middle schools. However, due to intense competition among students and their parents, as well as disparities in teacher quality, student achievement, and facilities among middle schools, the government abolished the middle school entrance exam in Seoul. After this policy was implemented, all students were randomly assigned to middle schools. Offices of Education in other provinces and metropolitan cities also followed suit by abolishing the entrance exams but later switched to preference-based assignments. Currently, only Seoul maintains random assignments based on school districts.

Seoul is divided into 11 school districts, each managed by a District Office of the Seoul Metropolitan Office of Education. Each district is comprised of one or two administrative districts in Seoul. For middle school assignment, each school district is further divided into three to six smaller sub-districts to account for commuting distances, resulting in a total of 46 sub-districts. The assignment process is based on a computerized lottery system that considers students' residential addresses at the time of elementary school graduation. This system does not differentiate between single-sex or coeducational schools or between private and public schools. Only a few specialized schools, such as those focusing on international education, arts, or sports, are allowed to select their students. Similarly, students who demonstrate special talents in sports or require special education can apply to specific schools.[8] Additionally, students who are disabled are assigned to schools near home. With these few exceptions, nearly all students are randomly assigned to schools. Table 2.1 shows the number of all-girl, all-boy, and coeducational middle schools in each district. In total, there are 44 all-girl, 48 all-boy, and 290 coeducational middle schools in Seoul.

[8]Students must win a sports competition or be approved by the elementary school principal to be acknowledged as having special talent in sports. Only after the District Office of the Seoul Metropolitan Office of Education approves their application can these students attend the school of their choice. There are fewer than 10 specialized middle schools in Seoul.

If a student and their family move into a different school district, they can request to be assigned to a school in the new district. However, this means the student will be assigned randomly within that school district, not to a specific school. Previous literature, such as Park et al. (2013), indicates that the proportion of students aged 14 to 16 who move to a new school district is small. Similarly, I show that the proportion of students whose middle school district differs from their elementary school district is 4.5 percent in my data. This alleviates concerns about families moving to be assigned to a particular middle school.

Private schools are not significantly different from public schools in the South Korean education system. Private schools constitute around 29 percent of all middle schools in Seoul. About 20 percent of private middle schools in South Korea are religious. A notable difference between private and public schools is that private schools have more flexible hiring practices, while the Seoul Metropolitan Office of Education assigns teachers to public schools. Public schools have been founded more recently than private schools due to governmental efforts to increase the number of schools.[9]

One limitation of this study is that it is hard to distinguish between private and single-sex schools. Approximately 86 percent of single-sex middle schools are private, whereas 10 percent of coeducational middle schools are private. Historically, schools founded early were predominantly single-sex, reflecting Confucian cultural norms that discouraged the mixing of adolescents from different sexes in the classroom. However, as societal norms evolved, later-established schools were not bound by this tradition. On average, public schools were established later than private schools. The Ministry of Education also mandated that some public single-sex middle schools transition to coeducational settings, while private

[9]Private schools in South Korea are defined as schools established and managed by corporations or individuals. When private schools were first established, they charged tuition from students and were considered more prestigious. However, as middle school education became compulsory, private schools are no longer allowed to collect tuition fees from students. For high schools, Hahn et al. (2018) shows that the share of teachers with M.A. degrees, the share of short-term and regular teachers, average compensation to teachers and staff, within-school base salary dispersion, and years since establishment differ between private and public schools.

single-sex schools were exempt from such directives. The main analysis does not control for private/public schools status. However, I also show the results from specifications that include a private school indicator in the robustness check.

2.3. Data

I use the Seoul Education Longitudinal Study 2010 (henceforth SELS 2010), a panel survey conducted by the Seoul Metropolitan Office of Education to evaluate the performance of schools and education policies in Seoul. The SELS 2010 sample consists of three cohorts of students in the fourth, seventh, and tenth grades in 2010. The survey sample was constructed using a two-stage stratified cluster sampling approach. In the first stage, schools were randomly sampled proportional to the number of schools within each school district. In the second stage, two classes from each grade cohort in the sampled schools were randomly selected. Sampled students were annually surveyed in July until they graduated high school.[10]

I restrict my sample to 5,059 students in the fourth-grade cohort, the only cohort who was surveyed in both elementary and middle school. I exclude students who did not participate in the survey during sixth grade, the final year of elementary school because survey responses from elementary school are essential for controlling baseline outcome measures and estimating the effects of being assigned to the different middle schools.[11] 3,485 students participated in the survey wave 6, when they are in ninth grade. I only use the surveys from the middle school years because high school placements are not random. I also drop the students whose gender information is missing. To estimate the effects of attending three years of single-sex schools, I further exclude students who transferred to different schools during their middle school years and those who did not respond to the survey in ninth grade.[12] My final sample

[10]The academic year begins in March in South Korea, so July marks the end of the first semester.
[11]Among 3,725 students who participated survey in wave 4 (i.e., when they were in seventh grade), 78 students, which is about 2.1 percent, did not participate in wave 3.
[12]96 students transferred during the middle school years, with 27 switching between single-sex and co-educational schools. There is almost no correlation between attending single-sex schools and transferring.

consists of 3,243 students from 289 schools. The number of schools in each district in my final sample is shown in Table 2.1. This includes 34 all-girl schools, 34 all-boy schools, and 221 coeducational middle schools.

The dataset includes surveys from students, parents, teachers, and school principals. Important individual and household characteristics available in SELS 2010 include birth order, family structure, parents' educational attainment, parents' employment status, and monthly household income. School-level characteristics in SELS 2010 are linked to publicly disclosed school information from the Korea Education & Research Information Service (KERIS). These measures include whether the school is private or public, the number of students and classes, the average teaching tenure, and the number of teachers by type for each year. Additionally, the school principal survey provides information on the principal's gender, educational attainment, and years of experience.

Table 2.2 presents summary statistics for individual and household characteristics of sixth-grade students. Columns 1 and 3 show the averages for students attending all-girl and all-boy schools, and Columns 2 and 4 show the averages for girls and boys in coeducational schools, respectively. In coeducational middle schools, 47.4 percent of students are female, similar to my entire sample. Across all school types, 44 to 47 percent of students are firstborn. Averages indicate that students in single-sex schools have a higher proportion of parents with college degrees, while students in coeducational schools are more likely to have parents who only graduated from high school.[13] Monthly household income, reported in thousands of

Attrition, which accounts for 203 observations lost during middle school, is also not correlated with single-sex school assignments.

[13]At the population level in 2012, 18 percent have attained less than upper secondary education, 41 percent have completed upper secondary or post-secondary non-tertiary education, and 42 percent have attained tertiary education among people aged 25 to 64 (OECD, 2014). Among people aged 25 to 34, 2 percent have attained less than upper secondary, 33 percent have completed upper secondary or post-secondary non-tertiary education, and 66 percent have attained tertiary education. My sample is similar to the younger age group.

South Korean Won (KRW) and adjusted to 2020 real values,[14] is slightly higher for students attending single-sex schools. Parental employment status and family structure are similar across the different school types.[15] The sample sizes for the schools are 316 students in all-girl schools, 283 students in all-boy schools, and 2,644 students in coeducational schools.

Table 2.5 shows summary statistics for school-level characteristics of ninth-grade students. Since school characteristics can vary yearly, I present the ninth-grade school characteristics, which are closely associated with students' outcomes in that grade. Columns 1, 2, and 3 represent all-girl, all-boy, and coeducational middle schools. The number of students per classroom, average teaching tenure, student-to-regular-teacher ratio, and student-to-regular-classroom-teacher ratio are similar across school types. Regular teachers, who are permanently employed by the Ministry of Education or the school, include regular classroom teachers, administrative teachers, principals, and vice principals. Regular classroom teachers primarily teach students. Non-regular teachers refer to those who are not permanently employed, such as part-time or contract-based teachers. There are fewer non-regular teachers per student in single-sex schools than in coeducational schools. Fully certified teachers hold a full teaching license (Level 1 or Level 2), complete a teacher education program, including coursework and student teaching, and pass the national teacher certification exam. All-girl and coeducational schools have more certified female teachers and fewer certified male teachers compared to all-boy schools. Principals in single-sex schools are also less likely to have graduate degrees and years of experience. Principals are more likely to be female in all-girl schools than in all-boy schools. Furthermore, 85.3 percent of all-girl and 88.2 percent of all-boy schools are private, while only 11.8 percent of coeducational schools are private. In

[14]Approximately 1,180 KRW is equivalent to 1 USD in 2020. In 2012, the average monthly household income was 4,294,118 KRW (in 2020 real values). Reported monthly household income is higher than the national average.

[15]At the population level for ages between 30 and 59, the male employment rate is around 90 percent, and the female employment rate is around 68 percent. I do not have data on the employment rate among married women. At the population level, one-parent families accounted for 10 percent in 2012, which is similar in my data. Source: Statistics Korea.

Table 2.18, I further divide public and private schools into all-girl, all-boy, and coeducational categories.

2.3.1. Outcome Measures. My outcomes of interest are cognitive test scores, self-esteem, self-directed learning approaches, goal-setting mindsets, social relationship, and school satisfaction. I use ninth-grade survey responses to estimate the effects of having attended single-sex schools for three years. These outcomes have been studied in the previous literature as they are related to later life outcomes. Self-esteem is positively associated with educational attainment or labor market outcomes (Heckman et al., 2006; de Araujo and Lagos, 2013). Social relationships reflect students' social and interpersonal skills. Social and interpersonal skills improve employment, wage, and occupation choices (Borghans et al., 2008c; Lindqvist and Vestman, 2011; Deming, 2017). Learning approaches can affect students' academic achievement and educational attainment, and goal-setting mindsets are related to students' motivation.

Korean, Mathematics, and English tests in SELS 2010 were developed for this survey. The tests cover the first-semester curriculum of each grade. I standardize test scores to mean zero and unit variance by grade by subject. Self-esteem, self-directed learning approaches, goal-setting mindsets, and social relationship are indices of several survey questions.[16] Student survey questionnaires ask students about self-esteem, self-directed learning approaches, goal-setting mindsets, social relationship, and school satisfaction. Students responded to each question using a 5-point Likert scale from 1 being 'strongly disagree' to 5 being 'strongly agree'. To create outcome indices, I standardize each survey response, take the average of survey responses, and standardize again to get a zero mean and unit variance.

The self-esteem index consists of five questions from the Rosenberg Self-Esteem Scale, a frequently used scale in Economics to measure global self-worth. The self-directed learning

[16]Cronbach's α, a commonly used reliability coefficient that assesses the internal consistency of survey items, ranges from 0.760 to 0.941. This indicates that survey items used for each index consistently measure the same construct.

approaches index measures whether students take initiative and motivation in their learning. This index includes nine questions developed from the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al., 1993). These questions cover various aspects of a student's learning process, such as learning effort (e.g., "I try my best to fully understand the school materials"), learning strategies (e.g., "I study important materials by summarizing key points or organizing them using charts, diagrams, or tables"), and learning attitudes (e.g., "I find studying enjoyable"). The goal-setting mindset index comprises five questions created based on the presence of and search for meaning in life (Steger et al., 2006). These questions include statements like "I have a clear goal that I want to achieve" and "I am working hard to achieve my goals." The social relationship index measures students' social and interpersonal skills with questions such as "Rather than being alone during breaks or lunchtime, I spend time with my friends," and "Even if I have arguments with my friends, we make up quickly." Lastly, school satisfaction is measured by a single question asking how satisfied students are with their school. All survey questions are listed in Table 2.20.[17]

2.3.2. Balancing Test. To verify the random assignment of children to schools, I perform a balancing test by regressing a single-sex school attendance variable on students' predetermined characteristics and baseline outcome measures in the sixth grade, controlling for elementary school fixed effects. Table 2.4 shows the results.[18] Most of the students' predetermined characteristics are not significantly different. Some baseline measures are different between students attending single-sex schools and those attending coeducational schools. Boys in single-sex schools are more likely to have higher Mathematics test scores and learning approaches than those in coeducational schools. Girls who are assigned to single-sex middle schools are more likely to be less satisfied with their elementary schools. However, F-statistics show that these differences are not jointly significant. Although these

[17]I show share of the missing outcome measures by school type in Table 2.16. I test whether missing outcome measures are correlated with attending single-sex schools. Attending single-sex schools is not correlated with missing outcome measures, as shown in Table 2.17.
[18]I show the balance in missing variables in Table 2.15.

are jointly not significant, one may still be worried about differences in baseline test scores and social-emotional skills. So my main specification controls all individual and household characteristics in the sixth grade.

2.4. Empirical Strategy

I begin by estimating the following baseline model:

$$y_{isp} = \alpha + \beta \cdot Singlesex_s + X_{ip} \cdot \gamma + Y_{ip} \cdot \delta + \lambda_p + e_{isp}$$

where y_{isp} is a summary measure of the outcome of an individual i who attend middle school s and attended elementary school p. Cognitive outcomes are Korean, Mathematics, and English test scores, and social-emotional outcomes include school satisfaction, social relationship, goal-setting mindsets, self-esteem, and learning approaches. $Singlesex_s$ equals one if the middle school s which student i attend is an all-girl or all-boy school. β, which is the main coefficient of interest, captures the impact of attending three years of single-sex middle school on students' test scores and social-emotional skills.

X_{ip} is a vector of individual-level or household-level characteristics at sixth grade. These include predetermined characteristics such as the gender of the student, a firstborn indicator, and parents' educational attainment, which is divided into three categories: less than high school, high school graduate, and college degree or above. X_{ip} also includes characteristics from sixth grade, the year before students attend middle schools, such as being in a two-parent family, the log of household income, and parents' employment. Additionally, I include Y_{ip}, a vector of baseline outcome measures in sixth grade. These measures are Korean, Mathematics, and English test scores; self-esteem; learning approaches; goal-setting mindsets; social relationship; and school satisfaction. I also show the regression coefficients without controls in Table 2.23 and Table 2.24.

I include elementary school fixed effects, λ_p, to control for the impact of elementary schools on students' skill development. Thus, my estimates capture the effects of attending single-sex middle schools compared to coeducational middle schools, given that students attended the same elementary school. These fixed effects also help absorb possible endogenous sorting across school districts. I do not use district fixed effects, as they may not adequately control for sorting across administrative districts (Sohn, 2016).[19] Students are randomly assigned to middle schools within their residential sub-districts inside the school district. Since there is no information about these sub-districts in the data, I control for elementary school instead. This approach effectively controls for possible sorting across school districts, as students are assigned to the nearest elementary schools. Finally, standard errors are clustered at the middle school level.

One may be concerned that among the 46 sub-districts, 20 sub-districts do not have all-girl middle schools and 15 sub-districts did not have all-boy middle schools as of 2020. Additionally, 10 sub-districts do not have either all-girl or all-boy middle schools. Ideally, this issue could be addressed by identifying whether students lived in sub-districts where all three school types were available. However, due to data limitations, I am unable to identify students' exact sub-districts. Instead, this concern is alleviated by including elementary school fixed effects, as the probability of being assigned to a single-sex school is constant for students coming from the same elementary school.[20] In addition, It is important to note that sub-districts were not disclosed to the public at that time, and students and parents were not aware of these sub-districts.

As 86 percent of single-sex schools are private, private school dummies are not included in my main specification due to a collinearity problem. Although assignments to private schools are random, there might be differences between private and public schools that affect students' skill development differently. To address this, I conduct a balancing test between

[19]A school district consists of one or two administrative districts.

[20]The results are also robust to restricting the sample to students in school districts where all three school types exist, with estimates similar to my main results.

students in private middle schools and those in public middle schools. The results, shown in Table 2.22, indicate that students attending private schools are similar to those attending public schools. Additionally, I demonstrate that the estimates controlling for private school are not significantly different from the estimates in my main specification.

2.5. Results

2.5.1. Test Scores. I begin by comparing the standardized test scores of ninth-grade students attending single-sex middle schools with those attending coeducational schools. Table 2.6 presents the results for test scores in Korean, Mathematics, and English. Columns 1, 3, and 5 display estimates that control for elementary school fixed effects, while Columns 2, 4, and 6 include both individual baseline controls and elementary school fixed effects. In the table, I also include mean values of each outcome measure among students attending coeducational schools. Panel A of Table 2.6 shows the results for the full sample, which includes both girls and boys. The estimates are positive and statistically significant, although they decrease slightly when individual baseline controls are added. To isolate the effects of attending single-sex schools from individual differences, I use the specification that controls for both elementary school fixed effects and individual baseline controls as my main model. The results suggest that attending three years of single-sex middle schools improves academic achievement across all three subjects.

Panels B and C show the results for girls and boys, respectively. Although the differences for girls are not statistically significant, those attending all-girl middle schools still perform better in all three subjects than their peers in coeducational schools. In contrast, boys attending all-boy schools have 0.236 standard deviation higher Mathematics scores and 0.319 standard deviation higher English scores in my main specification. The average test scores for boys are lower than those for girls, as shown in Table 2.13. Boys who attend all-boy schools perform better than average students in Mathematics and English. In summary,

attending single-sex schools appears to improve academic achievement for all students, with more gains for boys than for girls.

My estimates align with previous studies, which find larger effects of single-sex schools on boys. In the South Korean context, boys attending all-boy schools show overall improvements in test scores during middle school (Lee et al., 2014) and achieve higher scores on the CSAT, the exam required for college entrance (Park et al., 2013; Sohn, 2016). In a different country setting, Jackson (2019) also found that the increase in test scores was predominantly driven by boys. The impact of single-sex schools for girls, however, is mixed in the previous literature. Some studies find a positive association between attending all-girl high schools and higher CSAT scores (Park et al., 2013), while others, such as Sohn (2016), show negative effects on girls' CSAT scores after controlling for administrative district fixed effects and limiting the sample to public high school students. After controlling for elementary school fixed effects and individual baseline measures, the positive estimates across all three subjects suggest that girls attending single-sex middle schools may experience smaller improvements in academic achievement.

2.5.2. Social-Emotional Skills. I now explore whether single-sex schools are effective in improving social-emotional skills. Table 2.7 presents the results for self-esteem, learning approaches, goal-setting mindsets, social relationship, and school satisfaction. The odd-numbered columns show coefficients with only elementary school fixed effects, while the even-numbered columns include both individual baseline outcome measures and elementary school fixed effects. Panel A shows that students attending single-sex middle schools have 0.111 standard deviation higher self-esteem and 0.129 standard deviation higher learning approaches compared to those in coeducational schools. However, the estimates for goal-setting mindsets, social relationship, and school satisfaction are not statistically significant. Overall, single-sex schools seem more effective at enhancing self-esteem and learning approaches than coeducational schools.

Running separate regressions for girls and boys yields different results. As shown in Panel B, the estimates for girls are not statistically significant. However, girls in single-sex schools show increases in school satisfaction and decreases in social relationships, similar in magnitude to the self-esteem improvements observed in the entire sample. This suggests that girls in single-sex schools tend to have fewer social relationships but higher satisfaction with school compared to those in coeducational schools. Panel C reveals that boys in all-boy schools have 0.261 standard deviation higher self-esteem and 0.240 standard deviation higher learning approaches compared to boys in coeducational schools. They also show higher goal-setting mindsets, though these estimates are not statistically significantly different from zero. The improved learning approaches and potential increase in motivation, indicated by goal-setting mindsets, align with the higher test scores seen for boys in all-boy schools. In addition, girls and boys are not statistically significantly different in average ninth-grade self-esteem, learning approaches, goal-setting mindsets, and social relationship. In summary, the positive effects of single-sex schooling are driven by boys.

I compare my estimates to other studies using similar measures. For example, I compare the magnitude of the effects on learning approaches with teacher effects reported by Kraft (2019). The teacher effect on class effort is 0.183 standard deviations, while my estimate for boys' learning approaches, including effort and engagement, is 0.232 with individual controls. This suggests that attending all-boy schools instead of coeducational schools has an effect on learning approaches for boys similar to or larger than the teacher effect. Additionally, Gong et al. (2021) find that social acclimation and general satisfaction indices improve by 0.0767 standard deviations when the proportion of females in a classroom increases by 10 percent. This index correlates with social relationship and school satisfaction in this study, as social acclimation measures classmate interaction, and general satisfaction includes school life satisfaction and confidence in the future.

An all-girl school is analogous to having 100 percent girls in the classroom, while in coeducational schools, girls represent 47.4 percent of the student body. Thus, all-girl schools have a 52.6 percentage point larger share of girls, and all-boy schools have 47.4 percentage points fewer girls than coeducational schools. The estimates for the increase in girls' school satisfaction in all-girl schools and the decrease in boys' school satisfaction in all-boy schools are consistent with these findings, though smaller in magnitude. For example, a back-of-envelope calculation suggests that 47.4 percentage points fewer girls reduces social acclimation and general satisfaction index by 0.364 standard deviation, while my estimate is -0.063 for boys. Girls' school satisfaction also increases by 0.106 standard deviation when attending single-sex schools. Overall, these comparisons suggest that single-sex schooling has significant impacts on boys' social-emotional outcomes, comparable to both teacher and peer effects.

2.5.3. Robustness Checks. To assess the robustness of my results, I conduct several checks. First, I rerun the regressions including controls for private school status and other school-level characteristics, and test whether the effects of single-sex schools are affected. Additionally, I examine the effects of attending single-sex schools within samples of attending private and public schools separately to complement this test. Second, I exclude schools with fewer than 10 students in the sample and repeat the analyses to avoid skewing the results caused by a few observations representing the effect of one school. Lastly, I restrict the sample to students in districts that have all three types of schools: all-girl, all-boy, and coeducational schools.

Since most single-sex schools are private, I first investigate whether my results remain similar when a private school indicator and other school-level characteristics are included as controls in the regression. These school characteristics include the number of students per class in ninth grade, school-average teaching tenure in ninth grade, the number of regular and non-regular teachers per student in ninth grade, the female-to-male teacher ratio among level

1 and level 2 certified teachers, female principal indicator, and whether the principal holds a graduate degree. Results for test scores are presented in Table 2.8.[21] Columns 1, 4, and 7 present the baseline estimates, while Columns 2, 5, and 8 include the private school indicator, and Columns 3, 6, and 9 add both private school status and other school characteristics. The effects of single-sex schools on Mathematics test scores are influenced by private school status. Specifically, the effects of all-girl schools become larger and statistically significant, while the effects of all-boy schools shrink and lose significance. When additional school characteristics are controlled for, the effects of all-boy schools on Korean and Mathematics test scores reverse direction, and those on English test scores decrease and lose significance. Results for social-emotional skills, presented in Table 2.9, follow different patterns from test scores, with the effects largely robust to controlling for both the private school indicator and other school characteristics. In summary, while the estimates for test scores are sensitive to private school status and other school factors, the estimates for social-emotional skills remain stable. These findings suggest that the effects of single-sex schools are not entirely driven by their private status, particularly for social-emotional outcomes.

I further separate schools into private and public categories to explore the effects of single-sex schools within each group. Results are shown in Table 2.10. Students attending private coeducational schools are, on average, similar to those attending public coeducational schools. When comparing private and public coeducational schools by gender, girls attending public coeducational schools have higher Mathematics and English test scores than their counterparts in private coeducational schools. Boys attending private coeducational schools are not statistically significantly different from those attending public coeducational schools.

The effects of single-sex schools are not consistently stronger in private schools. In fact, private all-boy schools even show negative impacts on some outcomes, such as Korean and Mathematics test scores and social relationships. Conversely, public single-sex schools are

[21]I do the same analysis without using individual baseline controls. The results are shown in Tables 2.23 and 2.24. Estimates are a little greater, yet mostly similar.

effective in increasing girls' Mathematics and English test scores, as well as boys' learning approaches and social relationships. This suggests that the positive effects for boys are less likely to be driven by private school status.

Next, I restrict the sample in two ways and redo the analysis. First, I exclude schools with fewer than 10 students in the sample to avoid school effects being influenced by just a few observations. Due to the survey design, which tracks fourth-grade students from the first survey wave to middle school, the number of observations is not evenly distributed across middle schools, which can result in only a few students being assigned to certain schools. Limiting the sample to schools with at least 10 students removes 525 observations. Additionally, I restrict the analysis to districts that have all three types of schools: all-girl, all-boy, and coeducational middle schools. Some districts do not have single-sex schools, which leads to variations in gender composition in coeducational middle schools. To address this concern, I exclude two districts, South and Seongdong, removing 551 observations.

Tables 2.11 and 2.12 present the results. Compared to the baseline estimates in Columns 1, 4, and 7 of Table 2.11, Columns 2, 5, and 8 show that the effects on test scores remain relatively stable. Similarly, the estimates for social-emotional skills are comparable to the baseline, as shown in Columns 2, 5, 8, 11, and 14 of Table 2.12. Although the estimates for boys' Mathematics test scores and self-esteem lose statistical significance, they continue to indicate positive effects of attending single-sex schools. Furthermore, Columns 3, 6, and 9 of Table 2.11 and Columns 3, 6, 9, 12, and 15 of Table 2.12 show that the estimates excluding two districts are similar to the baseline. In summary, the single-sex school effects are not driven by a small number of observations per school or the absence of all three school types in certain districts.

2.6. Conclusion

In this paper, I examine the impact of single-sex schools on students' test scores in Korean, Mathematics, and English, as well as on their self-esteem, learning approaches, goal-setting

mindsets, social relationships, and school satisfaction in ninth grade. By controlling for elementary school fixed effects and baseline outcome measures in sixth grade, I isolate the effects of single-sex schools from elementary school effects and endogenous sorting across school districts.

The results indicate that boys experience significant improvements in Mathematics and English test scores after attending three years of all-boy schools. There is also suggestive evidence that attending all-girl schools enhances girls' test scores. Additionally, boys benefit from attending all-boy schools in terms of social-emotional skills, with improved self-esteem and better learning approaches. Although not statistically significant, boys attending all-boy schools tend to have clearer goals, greater motivation, and closer social relationships. In contrast, the effects on girls' social-emotional skills are mixed and noisier. In summary, all-boy schools are effective in enhancing both academic achievement and social-emotional skills, while all-girl schools may improve academic outcomes but do not necessarily enhance social-emotional skills. The differing effects on test scores and social-emotional skills for girls and boys in single-sex schools may lead to divergent skill sets, potentially affecting their performance in the labor market.

Tables

TABLE 2.1. Number of Middle Schools in Each District

	All-girl		All-boys		Coeducational		Total	
	Data	Actual	Data	Actual	Data	Actual	Data	Actual
Gangnam	5	5	3	5	18	29	26	39
Gangdong	1	5	4	5	25	34	30	44
Gangseo	3	3	4	4	22	33	29	40
South	0	0	0	2	21	31	21	33
East	4	6	2	2	16	21	22	29
Dongjak	1	2	1	3	21	27	23	32
North	1	1	0	1	30	37	31	39
West	6	7	5	7	29	32	40	46
Seongdong	1	1	0	0	21	22	22	23
Seongbuk	4	5	5	8	13	18	22	31
Center	8	9	10	11	5	6	23	26
Total	34	44	34	48	221	290	289	382

The table reports the number of all-girl, all-boy, and coeducational middle schools in each district. Odd columns show the observed number of schools in the data, while even columns report the actual number of schools in each district. Source: Korean Educational Development Institute (KEDI). Data for numbers of schools is from 2013.

TABLE 2.2. Descriptive Statistics of Individual and Household Characteristics

	Girls		Boys	
	All-girl	Coeducational	All-boy	Coeducational
Female	1.000	1.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Firstborn	0.449	0.468	0.470	0.451
	(0.498)	(0.499)	(0.500)	(0.498)
Age in 6th grade	11.923	11.918	11.926	11.939
	(0.312)	(0.307)	(0.297)	(0.305)
Mother's education				
Less than high school	0.013	0.019	0.021	0.018
	(0.112)	(0.137)	(0.144)	(0.133)
High school degree	0.297	0.403	0.336	0.363
	(0.458)	(0.491)	(0.473)	(0.481)
College degree	0.592	0.487	0.558	0.515
	(0.492)	(0.500)	(0.497)	(0.500)
Graduate degree	0.066	0.047	0.064	0.044
	(0.249)	(0.212)	(0.244)	(0.205)
Father's education				
Less than high school	0.006	0.014	0.011	0.014
	(0.079)	(0.116)	(0.103)	(0.116)
High school degree	0.244	0.301	0.244	0.264
	(0.430)	(0.459)	(0.430)	(0.441)
College degree	0.547	0.530	0.594	0.540
	(0.499)	(0.499)	(0.492)	(0.499)
Graduate degree	0.161	0.106	0.120	0.115
	(0.368)	(0.308)	(0.326)	(0.319)
Mother employed in 6th grade	0.525	0.540	0.502	0.531
	(0.500)	(0.499)	(0.501)	(0.499)
Father employed in 6th grade	0.940	0.933	0.943	0.936
	(0.238)	(0.250)	(0.231)	(0.245)
Monthly household income in 6th grade	5,458.362	5,185.759	5,370.729	5,273.346
	(3,505.953)	(4,935.958)	(2,966.144)	(4,187.193)
Two-parent family in 6th grade	0.899	0.895	0.894	0.887
	(0.302)	(0.306)	(0.308)	(0.316)
Observations	316	1251	283	1393

Standard deviation in parentheses. The monthly household income is in 1,000 South Korean Won(KRW) and transformed to 2020 real values. Approximately 1180 KRW is equivalent to 1 USD.

TABLE 2.3. Descriptive Statistics of Baseline Outcomes in Sixth Grade

	Girls		Boys	
	All-girl	Coeducational	All-boy	Coeducational
Korean test scores in 6th grade	0.275	0.177	-0.129	-0.190
	(0.836)	(0.899)	(1.017)	(1.069)
Math test scores in 6th grade	0.045	-0.018	0.122	-0.015
	(0.952)	(0.951)	(0.985)	(1.049)
English test scores in 6th grade	0.164	0.121	-0.064	-0.127
	(0.854)	(0.898)	(1.053)	(1.076)
Self-esteem in 6th grade	-0.070	0.001	0.170	-0.017
	(0.979)	(1.005)	(0.967)	(1.000)
Learning approaches in 6th grade	0.024	-0.017	0.216	-0.036
	(0.912)	(0.988)	(1.029)	(1.014)
Goal-setting mindsets in 6th grade	0.004	-0.033	0.207	-0.015
	(0.992)	(0.985)	(0.890)	(1.022)
Social relationship in 6th grade	0.011	-0.007	0.102	-0.023
	(0.990)	(0.968)	(0.982)	(1.036)
School satisfaction in 6th grade	-0.116	-0.038	0.016	0.062
	(0.941)	(0.950)	(1.140)	(1.019)
Observations	316	1251	283	1393

Standard deviation in parentheses.

TABLE 2.4. Balancing Test

	(1) Single-sex school		(2) All-girl school		(3) All-boy school	
	b	se	b	se	b	se
Firstborn	-0.006	0.027	-0.035	0.042	0.053	0.047
Age in 6th grade	-0.000	0.017	0.029	0.024	0.001	0.032
Log of household income in 6th grade	0.068	0.088	0.180	0.131	0.102	0.183
Two-parent family in 6th grade	-0.002	0.019	-0.011	0.023	-0.029	0.033
Mother's education						
Less than high school	-0.011*	0.006	-0.022*	0.013	0.001	0.004
High school degree	0.002	0.026	-0.038	0.040	0.046	0.042
College degree	0.013	0.027	0.054	0.043	-0.045	0.045
Graduate degree	0.002	0.012	0.013	0.020	0.007	0.026
Father's education						
Less than high school	-0.006	0.004	-0.013	0.011	-0.006	0.004
High school degree	-0.000	0.022	-0.041	0.034	0.043	0.039
College degree	0.007	0.026	-0.008	0.039	-0.007	0.045
Graduate degree	-0.003	0.020	0.046*	0.024	-0.017	0.036
Mother employed in 6th grade	0.005	0.032	0.068*	0.040	-0.047	0.046
Father employed in 6th grade	0.004	0.012	-0.021	0.017	0.030	0.021
Korean test score in 6th grade	0.075	0.067	0.110	0.075	0.096	0.106
Math test score in 6th grade	0.081	0.050	0.066	0.079	0.231**	0.094
English test score in 6th grade	0.035	0.051	0.017	0.073	0.163*	0.094
Self-esteem in 6th grade	0.012	0.063	-0.113	0.096	0.137	0.089
Learning approaches in 6th grade	0.087	0.054	0.018	0.071	0.241**	0.095
Goal-setting in 6th grade	0.075	0.057	0.003	0.083	0.149*	0.084
Social relationship in 6th grade	0.032	0.048	-0.029	0.063	0.030	0.104
School satisfaction in 6th grade	-0.104*	0.056	-0.139**	0.059	-0.154	0.110
Observations	3,173		1,536		1,637	
Test for joint significance:						
F-statistics	1.139		1.241		1.122	
p-value	0.307		0.223		0.328	
Primary school FE	Yes		Yes		Yes	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell presents a coefficient from a separate regression in which the dependent variable is a student's baseline characteristic and the independent variable is attending a single-sex middle school with primary school fixed effects. Standard errors, shown in parentheses, are clustered at the middle school level. F-statistics and p-value for joint significance are from a regression in which attending a single-sex middle school is regressed on all baseline characteristics with primary school fixed effects.

TABLE 2.5. School-Level Descriptive Statistics in Students' Ninth Grade

	All-girl	All-boy	Coeducational
Private	0.853	0.882	0.118
	(0.359)	(0.327)	(0.323)
Number of student per class	27.632	27.838	28.450
	(4.501)	(4.026)	(3.034)
Number of students in school	570.735	571.735	733.688
	(193.341)	(159.313)	(252.115)
Average teaching tenure	19.812	18.424	20.551
	(8.996)	(3.482)	(5.480)
Student-teacher ratio			
Regular teachers	16.381	17.296	17.022
	(3.382)	(3.028)	(2.542)
Regular classroom teachers	25.967	28.980	27.843
	(4.854)	(7.571)	(5.614)
Number of non-regular teachers	7.735	8.147	12.256
	(3.720)	(4.142)	(5.298)
Share of females among certified teachers	0.657	0.431	0.740
	(0.111)	(0.188)	(0.100)
Female principal	0.529	0.118	0.335
	(0.507)	(0.327)	(0.473)
Principal with graduate degree	0.471	0.559	0.805
	(0.507)	(0.504)	(0.397)
Experience of principal	11.303	8.000	13.075
	(14.611)	(12.738)	(15.347)
Observations	34	34	221

Standard deviation in parentheses. Regular teachers are teachers who are permanently employed by the Ministry of Education or the school, including regular classroom teachers, administrative teachers, principals, and vice principals. Regular classroom teachers primarily teach students. Non-regular teachers are teachers who are not permanently employed, such as part-time or contract-based teachers. Fully certified teachers hold a full teaching license (Level 1 or Level 2), complete a teacher education program, including coursework and student teaching, and pass the national teacher certification exam.

TABLE 2.6. Effects of Single-Sex Schooling on Test Scores in Ninth Grade

	Korean test scores		Math test scores		English test scores	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Entire Sample						
Single-sex school	0.228*	0.161	0.190**	0.149*	0.205**	0.148*
	(0.119)	(0.098)	(0.092)	(0.081)	(0.099)	(0.080)
Observations	3,233	3,108	3,234	3,109	3,228	3,103
Baseline mean	0.002	0.011	0.002	0.011	0.002	0.013
Panel B: Girls						
Single-sex school	0.235	0.197	0.106	0.057	0.118	0.090
	(0.146)	(0.133)	(0.130)	(0.112)	(0.116)	(0.098)
Observations	1,560	1,512	1,561	1,513	1,557	1,509
Baseline mean	0.197	0.200	-0.005	0.003	0.129	0.140
Panel C: Boys						
Single-sex school	0.157	0.097	0.295**	0.236**	0.406***	0.319***
	(0.128)	(0.116)	(0.127)	(0.114)	(0.115)	(0.106)
Observations	1,673	1,596	1,673	1,596	1,671	1,594
Baseline mean	-0.180	-0.169	0.009	0.018	-0.116	-0.108
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the effect of attending a single-sex school on Korean, math, and English test scores. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects. Column 1, 3, and 5 do not include individual-level controls, while column 2, 4, and 6 include individual-level controls in the regression. Individual-level controls include a female indicator, a firstborn indicator, log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.7. Effects of Single-Sex Schooling on Social-Emotional Skills in Ninth Grade

	Self-esteem		Learning approaches		Goal-setting		Friendship		School satisfaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Entire Sample										
Single-sex school	0.112*	0.111**	0.147**	0.129***	0.099	0.085	0.006	-0.007	0.007	0.042
	(0.066)	(0.052)	(0.059)	(0.045)	(0.070)	(0.061)	(0.069)	(0.060)	(0.069)	(0.070)
Observations	3,232	3,108	3,215	3,093	3,223	3,099	3,237	3,113	3,241	3,116
Baseline mean	0.001	0.003	-0.000	0.002	-0.001	0.000	-0.003	-0.006	0.002	0.003
Panel B: Girls										
Single-sex school	-0.034	0.036	0.038	0.073	-0.070	-0.007	-0.114	-0.070	0.064	0.106
	(0.096)	(0.077)	(0.098)	(0.079)	(0.112)	(0.104)	(0.088)	(0.081)	(0.094)	(0.090)
Observations	1,564	1,516	1,555	1,507	1,558	1,510	1,562	1,514	1,565	1,517
Baseline mean	-0.014	-0.012	-0.008	-0.009	-0.025	-0.022	-0.004	-0.009	-0.054	-0.058
Panel C: Boys										
Single-sex school	0.284**	0.261**	0.342***	0.240***	0.217**	0.158	0.093	0.061	-0.118	-0.063
	(0.113)	(0.110)	(0.099)	(0.081)	(0.109)	(0.096)	(0.128)	(0.117)	(0.109)	(0.113)
Observations	1,668	1,592	1,660	1,586	1,665	1,589	1,675	1,599	1,676	1,599
Baseline mean	0.015	0.017	0.007	0.012	0.023	0.022	-0.002	-0.004	0.054	0.061
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls		Yes		Yes		Yes		Yes		Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the effect of attending a single-sex school on students' self-esteem, learning approaches, goal-setting mindsets, social relationship, and school satisfaction. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects. Odd columns do not include individual-level controls, while even columns include individual-level controls in the regression. Individual-level controls include a female indicator, a firstborn indicator, log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, parents' employment status in 6th grade. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.8. Sensitivity of Single-Sex School Effects to School Characteristics: Test Scores

	Korean test scores			Math test scores			English test scores		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Entire Sample									
Single-sex school	0.161	0.185	0.210*	0.149*	0.209**	0.255***	0.148*	0.212**	0.234**
	(0.098)	(0.117)	(0.125)	(0.081)	(0.098)	(0.095)	(0.080)	(0.103)	(0.092)
Observations	3,108	3,108	3,018	3,109	3,109	3,019	3,103	3,103	3,014
Baseline mean	0.011	0.011	0.002	0.011	0.011	0.008	0.013	0.013	0.009
Panel B: Girls									
Single-sex school	0.197	0.267*	0.268*	0.057	0.301**	0.444***	0.090	0.181	0.182*
	(0.133)	(0.145)	(0.141)	(0.112)	(0.134)	(0.124)	(0.098)	(0.115)	(0.105)
Observations	1,512	1,512	1,454	1,513	1,513	1,455	1,509	1,509	1,452
Baseline mean	0.200	0.200	0.190	0.003	0.003	0.004	0.140	0.140	0.138
Panel C: Boys									
Single-sex school	0.097	0.043	-0.067	0.236**	0.081	0.028	0.319***	0.336**	0.209
	(0.116)	(0.138)	(0.161)	(0.114)	(0.136)	(0.148)	(0.106)	(0.133)	(0.139)
Observations	1,596	1,596	1,564	1,596	1,596	1,564	1,594	1,594	1,562
Baseline mean	-0.169	-0.169	-0.173	0.018	0.018	0.012	-0.108	-0.108	-0.112
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Private		Yes	Yes		Yes	Yes		Yes	Yes
School-level controls			Yes			Yes			Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the sensitivity of the results to the inclusion of a private school indicator and other school-level characteristics on Korean, math, and English test scores. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects and individual baseline controls. Individual-level controls include a female indicator, a firstborn indicator, a log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Column 2, 5, and 8 also include a private school indicator, and Column 3, 6, and 9 include a private school indicator and school-level characteristics in 9th grade in the regression. School-level controls, surveyed in 9th grade, include the number of students per classroom, average teaching experience, student-to-regular-teacher ratio, student-to-non-regular-teacher ratio, proportion of female teachers certified at level 1 and level 2, the gender of the school principal, and the educational attainment of the principal. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.9. Sensitivity of Single-Sex School Effects to School Characteristics: Social-Emotional Skills

	Self-esteem			Learning approaches			Goal-setting			Social relationship			School satisfaction		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Entire Sample															
Single-sex school	0.111**	0.097	0.107	0.129***	0.102	0.168**	0.085	0.039	0.014	-0.007	-0.004	-0.006	0.042	-0.065	-0.010
	(0.052)	(0.064)	(0.078)	(0.045)	(0.065)	(0.072)	(0.061)	(0.083)	(0.089)	(0.060)	(0.082)	(0.102)	(0.070)	(0.087)	(0.098)
Observations	3,108	3,108	3,016	3,093	3,093	3,002	3,099	3,099	3,008	3,113	3,113	3,021	3,116	3,116	3,024
Baseline mean	0.003	0.003	0.005	0.002	0.002	0.003	0.000	0.000	-0.004	-0.006	-0.006	-0.005	0.003	0.003	0.003
Panel B: Girls															
Single-sex school	0.036	0.038	-0.065	0.073	0.059	0.157	-0.007	-0.053	-0.116	-0.070	-0.083	-0.038	0.106	0.011	0.030
	(0.077)	(0.108)	(0.117)	(0.079)	(0.094)	(0.100)	(0.104)	(0.127)	(0.146)	(0.081)	(0.092)	(0.106)	(0.090)	(0.139)	(0.172)
Observations	1,516	1,516	1,456	1,507	1,507	1,448	1,510	1,510	1,451	1,514	1,514	1,454	1,517	1,517	1,457
Baseline mean	-0.012	-0.012	-0.013	-0.009	-0.009	-0.013	-0.022	-0.022	-0.030	-0.009	-0.009	-0.012	-0.058	-0.058	-0.058
Panel C: Boys															
Single-sex school	0.261**	0.268**	0.324**	0.240***	0.201**	0.364***	0.158	0.124	0.124	0.061	0.056	0.023	-0.063	-0.139	-0.149
	(0.110)	(0.123)	(0.141)	(0.081)	(0.101)	(0.122)	(0.096)	(0.119)	(0.133)	(0.117)	(0.146)	(0.179)	(0.113)	(0.120)	(0.150)
Observations	1,592	1,592	1,560	1,586	1,586	1,554	1,589	1,589	1,557	1,599	1,599	1,567	1,599	1,599	1,567
Baseline mean	0.017	0.017	0.022	0.012	0.012	0.018	0.022	0.022	0.020	-0.004	-0.004	0.001	0.061	0.061	0.060
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Private		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
School-level controls			Yes			Yes			Yes			Yes			Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the sensitivity of the results to the inclusion of a private school indicator and other school-level characteristics on self-esteem, learning approaches, goal-setting mindsets, social relationships, and school satisfaction. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects and individual baseline controls. Individual-level controls include a female indicator, a firstborn indicator, a log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Column 2, 5, and 8 also include a private school indicator, and Column 3, 6, and 9 include a private school indicator and school-level characteristics in 9th grade in the regression. School-level controls, surveyed in 9th grade, include the number of students per classroom, average teaching experience, student-to-regular-teacher ratio, student-to-non-regular-teacher ratio, proportion of female teachers certified at level 1 and level 2, the gender of the school principal, and the educational attainment of the principal. Baseline mean is each sample's average outcome index in the 6th grade.

Table 2.10. Private and Public Single-Sex School Comparison

	Korean		Math		English		Self-esteem		Learning		Goal-setting		Social relationship		School satisfaction	
	(1) Private	(2) Public	(3) Private	(4) Public	(5) Private	(6) Public	(7) Private	(8) Public	(9) Private	(10) Public	(11) Private	(12) Public	(13) Private	(14) Public	(15) Private	(16) Public
Panel A: Entire Sample																
Single-sex school	0.006	0.223	-0.017	0.271**	0.080	0.346***	0.068	-0.040	-0.009	0.234***	0.064	-0.107	-0.102	0.233	-0.001	-0.261
	(0.159)	(0.196)	(0.126)	(0.118)	(0.137)	(0.070)	(0.097)	(0.164)	(0.104)	(0.079)	(0.148)	(0.124)	(0.103)	(0.230)	(0.124)	(0.164)
Observations	759	2,349	759	2,350	757	2,346	763	2,345	758	2,335	756	2,343	761	2,352	762	2,354
Baseline mean	0.074	-0.010	0.053	-0.003	-0.018	0.023	0.045	-0.011	0.098	-0.029	0.091	-0.029	0.010	-0.011	0.011	0.001
Panel B: Girls																
Single-sex school	0.114	0.217	0.082	0.347*	0.151	0.250***	0.080	-0.058	-0.002	0.135	0.028	-0.091	-0.215	-0.182	0.160	-0.425
	(0.204)	(0.137)	(0.189)	(0.189)	(0.151)	(0.068)	(0.171)	(0.239)	(0.140)	(0.150)	(0.194)	(0.093)	(0.134)	(0.184)	(0.184)	(0.270)
Observations	397	1,115	397	1,116	396	1,113	399	1,117	395	1,112	396	1,114	396	1,118	397	1,120
Baseline mean	0.380	0.191	0.043	0.078	0.195	0.204	-0.048	0.002	0.015	0.022	0.027	-0.005	-0.025	0.036	0.016	-0.145
Panel C: Boys																
Single-sex school	-0.242	-0.290	-0.360**	0.020	0.015	0.478***	0.329*	0.120	0.147	0.247**	0.029	0.081	-0.281	0.815***	-0.425**	0.075
	(0.198)	(0.218)	(0.177)	(0.242)	(0.165)	(0.135)	(0.171)	(0.249)	(0.173)	(0.101)	(0.260)	(0.189)	(0.200)	(0.244)	(0.210)	(0.107)
Observations	362	1,234	362	1,234	361	1,233	364	1,228	363	1,223	360	1,229	365	1,234	365	1,234
Baseline mean	-0.223	-0.214	0.044	-0.073	-0.122	-0.186	0.129	-0.032	0.175	-0.075	0.098	-0.047	0.047	-0.040	0.111	0.110
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the effect of attending a single-sex school on Korean, math, and English test scores, self-esteem, learning approaches, goal-setting mindsets, social relationships, and school satisfaction. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects and individual-level controls. Individual-level controls include a female indicator, a firstborn indicator, a log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Baseline mean is each sample's average outcome index in the 6th grade. Among students attending private schools, 70.5 percent attend all-girl school among girls, and 66.8 percent attend all-boy school among boys. Among students attending public schools, 2.1 percent attend all-girl schools among girls, and 2.2 percent attend all-boy schools among boys.

TABLE 2.11. Robustness of Single-Sex School Effects to Different Sample Restriction: Test Scores

	Korean test scores			Math test scores			English test scores		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Entire Sample									
Single-sex school	0.161	0.121	0.159	0.149*	0.163	0.118	0.148*	0.096	0.132
	(0.098)	(0.143)	(0.105)	(0.081)	(0.117)	(0.086)	(0.080)	(0.119)	(0.086)
Observations	3,108	2,609	2,565	3,109	2,609	2,566	3,103	2,604	2,564
Baseline mean	0.011	0.029	0.007	0.011	0.021	0.015	0.013	0.019	0.015
Panel B: Girls									
Single-sex school	0.197	0.284	0.239*	0.057	0.134	0.040	0.090	0.083	0.103
	(0.133)	(0.181)	(0.144)	(0.112)	(0.156)	(0.123)	(0.098)	(0.139)	(0.107)
Observations	1,512	1,297	1,252	1,513	1,297	1,253	1,509	1,294	1,252
Baseline mean	0.200	0.224	0.213	0.003	0.030	0.021	0.140	0.153	0.153
Panel C: Boys									
Single-sex school	0.097	-0.050	0.050	0.236**	0.160	0.205*	0.319***	0.302*	0.262**
	(0.116)	(0.183)	(0.117)	(0.114)	(0.204)	(0.119)	(0.106)	(0.172)	(0.111)
Observations	1,596	1,312	1,313	1,596	1,312	1,313	1,594	1,310	1,312
Baseline mean	-0.169	-0.164	-0.190	0.018	0.013	0.010	-0.108	-0.113	-0.118
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School obs $\geq$ 10		Yes			Yes			Yes	
Districts with all school types			Yes			Yes			Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column 1, 4, and 7 present the main results; Column 2, 5, and 8 present the results of students attending schools that have observations greater than 9; Column 3, 6, and 9 present the results of students attending schools in school districts that have all-girls, all-boys, and coeducational schools. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects and individual-level controls. Individual-level controls include a female indicator, a firstborn indicator, a log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.12. Robustness of Single-Sex School Effects to Different Sample Restriction: Social-Emotional Skills

	Self-esteem			Learning approaches			Goal-setting			Social relationship			School satisfaction		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Entire Sample															
Single-sex school	0.111**	0.095	0.129**	0.129***	0.167***	0.147***	0.085	0.084	0.084	-0.007	0.021	-0.009	0.042	0.077	0.019
	(0.052)	(0.063)	(0.054)	(0.045)	(0.056)	(0.046)	(0.061)	(0.077)	(0.066)	(0.060)	(0.076)	(0.062)	(0.070)	(0.092)	(0.074)
Observations	3,108	2,600	2,567	3,093	2,588	2,556	3,099	2,590	2,562	3,113	2,605	2,570	3,116	2,608	2,573
Baseline mean	0.003	0.019	0.016	0.002	0.016	0.016	0.000	0.002	0.004	-0.006	-0.016	0.011	0.003	0.007	-0.002
Panel B: Girls															
Single-sex school	0.036	0.026	0.040	0.073	0.125	0.083	-0.007	-0.023	-0.023	-0.070	-0.038	-0.076	0.106	0.155	0.052
	(0.077)	(0.088)	(0.085)	(0.079)	(0.118)	(0.085)	(0.104)	(0.151)	(0.113)	(0.081)	(0.120)	(0.083)	(0.090)	(0.116)	(0.097)
Observations	1,516	1,294	1,256	1,507	1,285	1,249	1,510	1,288	1,252	1,514	1,293	1,254	1,517	1,296	1,257
Baseline mean	-0.012	0.015	-0.008	-0.009	0.019	-0.007	-0.022	-0.004	-0.027	-0.009	-0.001	-0.001	-0.058	-0.044	-0.071
Panel C: Boys															
Single-sex school	0.261**	0.136	0.289**	0.240***	0.238**	0.292***	0.158	0.158	0.177*	0.061	0.044	0.090	-0.063	-0.148	-0.096
	(0.110)	(0.164)	(0.112)	(0.081)	(0.110)	(0.083)	(0.096)	(0.112)	(0.103)	(0.117)	(0.207)	(0.123)	(0.113)	(0.172)	(0.119)
Observations	1,592	1,306	1,311	1,586	1,303	1,307	1,589	1,302	1,310	1,599	1,312	1,316	1,599	1,312	1,316
Baseline mean	0.017	0.022	0.038	0.012	0.014	0.038	0.022	0.009	0.033	-0.004	-0.030	0.023	0.061	0.059	0.065
Elementary school FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School obs ≥ 10		Yes			Yes			Yes			Yes			Yes	
Districts with all school types			Yes			Yes			Yes			Yes			Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column 1, 4, 7, 10, and 13 present the main results; Column 2, 5, 8, 11, and 14 present the results of students attending schools that have observations greater than 9; Column 3, 6, 9, 12, and 15 present the results of students attending schools in school districts that have all-girls, all-boys, and coeducational schools. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. All regressions include primary school fixed effects and individual-level controls. Individual-level controls include a female indicator, a firstborn indicator, a log of household income in 6th grade, family structure in 6th grade, parents' educational attainment, and parents' employment status in 6th grade. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.13. Descriptive Statistics of Outcome Measures in Ninth Grade

	Girls	Boys	Differences
Korean test scores in 9th grade	0.235	-0.214	-0.451***
	(0.925)	(1.015)	(0.034)
Math test scores in 9th grade	0.062	-0.049	-0.118***
	(0.972)	(1.019)	(0.035)
English test scores in 9th grade	0.193	-0.171	-0.367***
	(0.945)	(1.012)	(0.035)
Self-esteem in 9th grade	-0.015	0.018	0.036
	(0.986)	(1.011)	(0.035)
Learning approaches in 9th grade	0.019	-0.014	-0.031
	(0.962)	(1.035)	(0.036)
Goal-setting in 9th grade	0.004	-0.005	-0.012
	(0.968)	(1.021)	(0.035)
Social relationship in 9th grade	0.022	-0.016	-0.034
	(0.958)	(1.038)	(0.036)
School satisfaction in 9th grade	-0.100	0.107	0.203***
	(0.942)	(1.033)	(0.035)
Observations	1,567	1,676	3,171

Standard deviations in parentheses. *** $p < 0.01$.

TABLE 2.14. Descriptive Statistics of Missing Individual Baseline Controls

	Girls		Boys	
	All-girls	Coeducational	All-boys	Coeducational
Missing firstborn	0.019	0.027	0.018	0.047
	(0.137)	(0.163)	(0.132)	(0.211)
Missing household income in 6th grade	0.022	0.030	0.035	0.029
	(0.147)	(0.169)	(0.185)	(0.169)
Missing mother's education	0.032	0.044	0.021	0.060
	(0.175)	(0.205)	(0.144)	(0.238)
Missing father's education	0.041	0.050	0.032	0.067
	(0.199)	(0.217)	(0.176)	(0.251)
Missing parents' employment status in 6th grade	0.009	0.006	0.007	0.008
	(0.097)	(0.080)	(0.084)	(0.089)
Missing family structure in 6th grade	0.000	0.003	0.004	0.006
	(0.000)	(0.056)	(0.059)	(0.080)
Missing Korean test scores in 6th grade	0.006	0.002	0.000	0.001
	(0.079)	(0.040)	(0.000)	(0.038)
Missing math test scores in 6th grade	0.006	0.002	0.000	0.001
	(0.079)	(0.040)	(0.000)	(0.038)
Missing English test scores in 6th grade	0.006	0.002	0.000	0.001
	(0.079)	(0.040)	(0.000)	(0.038)
Missing self-esteem in 6th grade	0.009	0.002	0.004	0.010
	(0.097)	(0.049)	(0.059)	(0.100)
Missing learning approaches in 6th grade	0.019	0.010	0.014	0.021
	(0.137)	(0.098)	(0.118)	(0.143)
Missing goal-setting mindsets in 6th grade	0.009	0.008	0.014	0.012
	(0.097)	(0.089)	(0.118)	(0.110)
Missing social relationship in 6th grade	0.003	0.006	0.007	0.008
	(0.056)	(0.075)	(0.084)	(0.089)
Missing school satisfaction in 6th grade	0.006	0.005	0.011	0.006
	(0.079)	(0.069)	(0.103)	(0.076)
Observations	316	1251	283	1393

Standard deviation in parentheses. This table presents descriptive statistics of missing individual baseline controls by gender in single-sex schools and education schools.

TABLE 2.15. Balancing Test in Missing Individual Baseline Controls

	(1) Single-sex school		(2) All-girl school		(3) All-boy school	
	b	se	b	se	b	se
Missing firstborn	0.003	0.007	0.014	0.008	-0.007	0.013
Missing household income in 6th grade	-0.007	0.010	-0.026*	0.015	-0.002	0.021
Missing family structure in 6th grade	-0.006	0.004	-0.006	0.005	-0.005	0.005
Missing mother's education	-0.008	0.008	-0.005	0.013	-0.015	0.015
Missing father's education	0.001	0.010	0.018	0.013	-0.020	0.018
Missing parents' employment status in 6th grade	-0.006	0.007	-0.002	0.013	-0.000	0.004
Missing Korean test score in 6th grade	0.002	0.002	0.000	0.003	0.000	.
Missing math test score in 6th grade	0.002	0.002	0.000	0.003	0.000	.
Missing English test score in 6th grade	0.002	0.002	0.000	0.003	0.000	.
Missing self-esteem in 6th grade	0.000	0.004	0.005	0.004	0.001	0.007
Missing learning approaches in 6th grade	-0.002	0.005	0.008	0.005	-0.017	0.013
Missing goal-setting in 6th grade	-0.006	0.006	-0.001	0.007	-0.012	0.012
Missing social relationship in 6th grade	0.002	0.004	0.004	0.004	0.005	0.007
Missing school satisfaction in 6th grade	0.010*	0.005	0.007	0.008	0.013	0.008
Observations	3,243		1,567		1,676	
Primary school FE	Yes		Yes		Yes	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell presents a coefficient from a separate regression in which the dependent variable is a missing indicator of individual baseline controls and the independent variable is attending a single-sex middle school with primary school fixed effects. Standard errors, shown in parentheses, are clustered at the middle school level.

TABLE 2.16. Descriptive Statistics of Missing Outcome Measures

	Girls		Boys	
	All-girls	Coeducational	All-boys	Coeducational
Missing Korean test scores in 9th grade	0.013	0.002	0.011	0.000
	(0.112)	(0.049)	(0.103)	(0.000)
Missing math test scores in 9th grade	0.009	0.002	0.011	0.000
	(0.097)	(0.049)	(0.103)	(0.000)
Missing English test scores in 9th grade	0.013	0.005	0.014	0.001
	(0.112)	(0.069)	(0.118)	(0.027)
Missing self-esteem in 9th grade	0.000	0.002	0.007	0.004
	(0.000)	(0.049)	(0.084)	(0.066)
Missing learning approaches in 9th grade	0.009	0.007	0.018	0.008
	(0.097)	(0.085)	(0.132)	(0.089)
Missing goal-setting mindsets in 9th grade	0.006	0.006	0.018	0.004
	(0.079)	(0.075)	(0.132)	(0.066)
Missing social relationship in 9th grade	0.006	0.002	0.000	0.001
	(0.079)	(0.049)	(0.000)	(0.027)
Missing school satisfaction in 9th grade	0.006	0.000	0.000	0.000
	(0.079)	(0.000)	(0.000)	(0.000)
Observations	316	1251	283	1393

Standard deviation in parentheses. This table presents descriptive statistics of missing outcome measures in 9th grade by gender.

TABLE 2.17. Balance in Missing Outcomes

	(1) Single-sex school		(2) All-girl school		(3) All-boy school	
	b	se	b	se	b	se
Missing Korean test score in 9th grade	0.011	0.009	0.005	0.011	0.025	0.022
Missing math test score in 9th grade	0.008	0.009	-0.002	0.010	0.025	0.022
Missing English test score in 9th grade	0.010	0.010	0.001	0.010	0.026	0.022
Missing self-esteem in 9th grade	0.000	0.003	-0.000	.	0.008	0.007
Missing learning approaches in 9th grade	0.003	0.007	-0.009	0.008	0.016	0.017
Missing goal-setting in 9th grade	0.004	0.004	-0.001	0.006	0.009	0.008
Missing social relationship in 9th grade	0.005	0.004	0.009	0.007	0.000	.
Missing school satisfaction in 9th grade	0.005*	0.002	0.005	0.003	0.000	.
Observations	3,243		1,567		1,676	
Primary school FE	Yes		Yes		Yes	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell presents a coefficient from a separate regression in which the dependent variable is a missing indicator of outcome measures and the independent variable is attending a single-sex middle school with primary school fixed effects. Standard errors, shown in parentheses, are clustered at the middle school level.

TABLE 2.18. School-Level Descriptive Statistics in Students' Ninth Grade

	All-girls		All-boys		Coeducational	
	Public	Private	Public	Private	Public	Private
Number of student per class	24.627	28.150	27.840	27.838	28.292	29.762
	(4.138)	(4.420)	(3.211)	(4.169)	(3.047)	(2.626)
Number of students in school	447.800	591.931	648.250	561.533	739.344	686.478
	(152.431)	(193.836)	(174.931)	(157.482)	(259.576)	(175.214)
Average teaching tenure	18.200	20.111	16.667	18.600	20.771	18.792
	(4.817)	(9.609)	(3.512)	(3.490)	(5.603)	(4.054)
Student-teacher ratio						
Regular teachers	12.720	17.012	17.123	17.319	16.672	19.946
	(2.452)	(3.133)	(1.977)	(3.166)	(2.145)	(3.595)
Regular classroom teachers	24.105	26.288	29.558	28.903	27.176	33.419
	(5.131)	(4.826)	(3.781)	(7.981)	(3.797)	(12.042)
Number of non-regular teachers	8.200	7.655	13.000	7.500	12.339	11.565
	(3.347)	(3.829)	(3.367)	(3.830)	(5.351)	(4.888)
Share of females among certified teachers	0.777	0.636	0.749	0.389	0.764	0.544
	(0.106)	(0.099)	(0.045)	(0.155)	(0.065)	(0.126)
Female principal	0.600	0.517	0.750	0.033	0.369	0.077
	(0.548)	(0.509)	(0.500)	(0.183)	(0.484)	(0.272)
Principal with graduate degree	1.000	0.379	1.000	0.500	0.846	0.500
	(0.000)	(0.494)	(0.000)	(0.509)	(0.362)	(0.510)
Experience of principal	8.000	11.893	1.500	8.897	12.853	14.913
	(14.680)	(14.788)	(2.380)	(13.340)	(15.372)	(15.353)
Observations	5	29	4	30	195	26

Standard deviation in parentheses. This table presents school-level characteristics in 9th grade in all-girls, all-boys, and coeducational schools. Column 1, 3, and 5 present public schools in each school type, and Column 2, 4, and 6 present private schools in each school type.

TABLE 2.19. Comparison of School Characteristics in Students' Ninth Grade

	(1) Single-sex school	(2) All-girl school	(3) All-boy school
Private school	0.781***	0.753***	0.761***
	(0.049)	(0.063)	(0.067)
Number of student per class	0.174	0.044	0.421
	(0.587)	(0.806)	(0.754)
Number of students in school	-108.680***	-110.082***	-104.800***
	(29.860)	(36.982)	(37.245)
Average teaching tenure	0.797	1.809	-0.060
	(1.032)	(1.784)	(0.774)
Regular teachers	0.260	-0.126	0.692
	(0.446)	(0.524)	(0.642)
Regular classroom teachers	-1.131	-2.186**	0.255
	(0.869)	(0.941)	(1.271)
Number of non-regular teachers	-4.831***	-5.122***	-4.289***
	(0.684)	(0.800)	(0.927)
Share of females among certified teachers	-0.212***	-0.096***	-0.325***
	(0.027)	(0.023)	(0.036)
Female principal	-0.001	0.227**	-0.206***
	(0.072)	(0.101)	(0.069)
Principal with graduate degree	-0.428***	-0.461***	-0.372***
	(0.069)	(0.093)	(0.090)
Experience of principal	-1.167	0.091	-2.876
	(2.290)	(3.077)	(2.733)
Observations	279	246	246

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell presents a coefficient from a separate regression in which the dependent variable is a school-level characteristic, and the independent variable is a single-sex school indicator.

TABLE 2.20. SELS 2010 Survey Questionnaires

Indices	Questions
Self-esteem: (Rosenberg Self-Esteem scale)	I feel that I have a number of good qualities. I am able to do things as well as most other people. I feel that I'm a person of worth. I take a positive attitude toward myself. On the whole, I am satisfied with myself.
Social relationships:	I have a friend whom I can trust and talk to. Rather than being alone during breaks or lunchtime, I spend time with my friends. Even if I have arguments with my friends, we make up quickly. I help my friends who need assistance.
Learning approaches:	I think about how to connect newly learned information with what I already know. I study important materials by summarizing key points or organizing them in tables or mind maps. I check myself to see if I understand the class materials well. I try my best to fully understand the school materials. I try to stick to the study schedule I planned as much as possible. If there is anything I don't understand while studying or doing assignments, I look it up in books or online. I believe I will eventually understand any difficult material. I find studying enjoyable. I can handle my own work well without anyone telling me what to do.
Goal-setting mindsets:	I have a clear goal that I want to achieve. I know what I need to do to achieve my goals. I am working hard to achieve my goals. The study I am doing now will help me achieve my future goals. If I achieve my future goals, I believe I can also contribute to society.

TABLE 2.21. Reliability Measures: Cronbach's α

	Primary school	Middle school
Self-esteem	0.941	0.941
Friendship	0.760	0.782
Learning strategies	0.851	0.835
Learning efforts	0.816	0.823
Learning attitudes	0.780	0.776
Goal-setting mindsets	0.889	0.902

Cronbach's α is measured in the fourth survey wave.

TABLE 2.22. Balancing Test Between Private School and Public School Students

	(1) Entire sample		(2) Girls		(3) Boys	
	b	se	b	se	b	se
Firstborn	0.023	0.025	-0.028	0.040	0.112***	0.041
Age in 6th grade	-0.026*	0.015	-0.018	0.023	-0.010	0.025
Log of household income in 6th grade	-0.044	0.037	-0.070	0.066	0.002	0.045
Two-parent family in 6th grade	-0.000	0.016	-0.005	0.021	-0.012	0.028
Mother's education						
Less than high school	0.004	0.005	0.009	0.007	0.002	0.006
High school degree	-0.010	0.027	-0.027	0.043	0.015	0.040
College degree	0.010	0.027	0.006	0.045	-0.019	0.044
Graduate degree	-0.003	0.010	0.011	0.013	0.001	0.019
Father's education						
Less than high school	0.003	0.004	0.003	0.008	-0.001	0.005
High school degree	-0.019	0.019	-0.030	0.032	-0.013	0.034
College degree	0.032	0.022	-0.005	0.036	0.028	0.041
Graduate degree	-0.016	0.017	0.032	0.023	-0.014	0.029
Mother employed in 6th grade	-0.016	0.026	0.024	0.037	-0.036	0.039
Father employed in 6th grade	-0.007	0.010	-0.023	0.015	0.010	0.019
Korean test score in 6th grade	0.032	0.057	-0.034	0.068	0.123	0.084
Math test score in 6th grade	0.069	0.047	0.010	0.080	0.170**	0.077
English test score in 6th grade	-0.019	0.049	-0.119	0.082	0.100	0.077
Self-esteem in 6th grade	0.029	0.050	-0.024	0.082	0.004	0.068
Learning approaches in 6th grade	0.094*	0.048	0.002	0.072	0.186***	0.071
Goal-setting in 6th grade	0.046	0.048	0.022	0.078	0.080	0.069
Social relationship in 6th grade	-0.002	0.041	-0.061	0.060	-0.022	0.074
School satisfaction in 6th grade	-0.022	0.047	-0.063	0.055	-0.100	0.085
Test for joint significance:						
F-statistics	0.820		1.008		1.116	
p-value	0.688		0.454		0.334	
Observations	3,224		1,559		1,665	
Primary school FE	Yes		Yes		Yes	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell presents a coefficient from a separate regression in which the dependent variable is a student's baseline characteristic and the independent variable is attending a private middle school with primary school fixed effects. Sample is defined by column. Standard errors, shown in parentheses, are clustered at the middle school level.

TABLE 2.23. Sensitivity to the Inclusion of School Characteristics Without Individual Controls: Test Scores

	Korean test scores				Math test scores				English test scores			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Entire Sample												
Single-sex school	0.195	0.228*	0.250*	0.335**	0.144	0.190**	0.249**	0.370***	0.152	0.205**	0.280**	0.381***
	(0.126)	(0.119)	(0.146)	(0.159)	(0.111)	(0.092)	(0.121)	(0.116)	(0.123)	(0.099)	(0.131)	(0.131)
Observations	3,233	3,233	3,233	3,141	3,234	3,234	3,234	3,142	3,228	3,228	3,228	3,137
Baseline mean	0.002	0.002	0.002	-0.007	0.002	0.002	0.002	-0.001	0.002	0.002	0.002	-0.002
Panel B: Girls												
Single-sex school	0.346***	0.235	0.384**	0.509***	0.127	0.106	0.389**	0.640***	0.119	0.118	0.298*	0.481***
	(0.118)	(0.146)	(0.168)	(0.178)	(0.148)	(0.130)	(0.169)	(0.165)	(0.145)	(0.116)	(0.153)	(0.152)
Observations	1,560	1,560	1,560	1,501	1,561	1,561	1,561	1,502	1,557	1,557	1,557	1,499
Baseline mean	0.197	0.197	0.197	0.188	-0.005	-0.005	-0.005	-0.005	0.129	0.129	0.129	0.126
Panel C: Boys												
Single-sex school	-0.016	0.157	0.068	-0.035	0.150	0.295**	0.129	0.124	0.145	0.406***	0.403**	0.303*
	(0.163)	(0.128)	(0.160)	(0.182)	(0.157)	(0.127)	(0.154)	(0.167)	(0.177)	(0.115)	(0.157)	(0.163)
Observations	1,673	1,673	1,673	1,640	1,673	1,673	1,673	1,640	1,671	1,671	1,671	1,638
Baseline mean	-0.180	-0.180	-0.180	-0.185	0.009	0.009	0.009	0.002	-0.116	-0.116	-0.116	-0.120
Elementary school FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
Individual controls												
Private			Yes	Yes			Yes	Yes			Yes	Yes
School-level controls				Yes				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the sensitivity of the results to the inclusion of a private school indicator and other school-level characteristics on self-esteem, learning approaches, goal-setting mindsets, social relationships, and school satisfaction. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. Columns 1, 5, and 9 show results without primary school fixed effects and individual controls. Columns 2, 6, and 10 present results with primary school fixed effects without individual controls. Columns 3, 7, and 11 also present results with a private school indicator, and Columns 4, 8, and 12 present results with a private school indicator and school-level characteristics in 9th grade in the regression. School-level controls, surveyed in 9th grade, include the number of students per classroom, average teaching experience, student-to-regular-teacher ratio, student-to-non-regular-teacher ratio, proportion of female teachers certified at level 1 and level 2, the gender of the school principal, and the educational attainment of the principal. Baseline mean is each sample's average outcome index in the 6th grade.

TABLE 2.24. Sensitivity to the Inclusion of School Characteristics Without Individual Controls: Social-emotional Skills

	Self-esteem				Learning approaches				Goal-setting				Social relationship				School satisfaction			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel A: Entire Sample																				
Single-sex school	0.097*	0.112*	0.093	0.117	0.173***	0.147**	0.109	0.218**	0.097	0.099	0.029	0.019	0.035	0.006	0.014	0.007	0.030	0.007	-0.127	-0.100
	(0.059)	(0.066)	(0.078)	(0.092)	(0.063)	(0.059)	(0.086)	(0.099)	(0.062)	(0.070)	(0.088)	(0.094)	(0.051)	(0.069)	(0.091)	(0.104)	(0.070)	(0.069)	(0.093)	(0.104)
Observations	3,232	3,232	3,232	3,138	3,215	3,215	3,215	3,122	3,223	3,223	3,223	3,130	3,237	3,237	3,237	3,143	3,241	3,241	3,241	3,147
Baseline mean	0.001	0.001	0.001	0.003	-0.000	-0.000	-0.000	0.001	-0.001	-0.001	-0.001	-0.005	-0.003	-0.003	-0.003	-0.002	0.002	0.002	0.002	0.002
Panel B: Girls																				
Single-sex school	-0.044	-0.034	-0.026	-0.109	0.057	0.038	0.052	0.210*	-0.012	-0.070	-0.148	-0.204	-0.078	-0.114	-0.107	-0.081	0.179**	0.064	-0.051	-0.043
	(0.078)	(0.096)	(0.127)	(0.147)	(0.084)	(0.098)	(0.121)	(0.126)	(0.084)	(0.112)	(0.123)	(0.141)	(0.072)	(0.088)	(0.105)	(0.113)	(0.090)	(0.094)	(0.139)	(0.177)
Observations	1,564	1,564	1,564	1,503	1,555	1,555	1,555	1,495	1,558	1,558	1,558	1,498	1,562	1,562	1,562	1,501	1,565	1,565	1,565	1,504
Baseline mean	-0.014	-0.014	-0.014	-0.016	-0.008	-0.008	-0.008	-0.013	-0.025	-0.025	-0.025	-0.033	-0.004	-0.004	-0.004	-0.007	-0.054	-0.054	-0.054	-0.053
Panel C: Boys																				
Single-sex school	0.254***	0.284**	0.312**	0.393***	0.296***	0.342***	0.286**	0.469***	0.216**	0.217**	0.179	0.230	0.151**	0.093	0.115	0.104	-0.105	-0.118	-0.184	-0.209
	(0.081)	(0.113)	(0.125)	(0.137)	(0.090)	(0.099)	(0.121)	(0.135)	(0.087)	(0.109)	(0.134)	(0.143)	(0.065)	(0.128)	(0.162)	(0.192)	(0.106)	(0.109)	(0.126)	(0.153)
Observations	1,668	1,668	1,668	1,635	1,660	1,660	1,660	1,627	1,665	1,665	1,665	1,632	1,675	1,675	1,675	1,642	1,676	1,676	1,676	1,643
Baseline mean	0.015	0.015	0.015	0.021	0.007	0.007	0.007	0.013	0.023	0.023	0.023	0.021	-0.002	-0.002	-0.002	0.002	0.054	0.054	0.054	0.053
Elementary school FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
Individual controls																				
Private			Yes	Yes			Yes	Yes			Yes	Yes			Yes	Yes			Yes	Yes
School-level controls				Yes				Yes				Yes				Yes				Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table presents the sensitivity of the results to the inclusion of a private school indicator and other school-level characteristics on self-esteem, learning approaches, goal-setting mindsets, social relationships, and school satisfaction. Each cell reports a result from a separate regression. Standard errors, shown in parentheses, are clustered at the middle school level. Columns 1, 5, 9, 13, and 17 show results without primary school fixed effects and individual controls. Columns 2, 6, 10, 14, and 18 present results with primary school fixed effects without individual controls. Columns 3, 7, 11, 15, and 19 also present results with a private school indicator, and Columns 4, 8, 12, 16, and 20 present results with a private school indicator and school-level characteristics in 9th grade in the regression. School-level controls, surveyed in 9th grade, include the number of students per classroom, average teaching experience, student-to-regular-teacher ratio, student-to-non-regular-teacher ratio, proportion of female teachers certified at level 1 and level 2, the gender of the school principal, and the educational attainment of the principal. Baseline mean is each sample's average outcome index in the 6th grade.

CHAPTER 3

School Entry Cutoff and Adolescent Mental Health: Evidence from a South Korean Reform (with Baiyu Zhou)

3.1. Introduction

Childhood through adolescence is a critical period of development marked by significant physical, emotional, and cognitive changes. Mental health during this time shapes not only immediate experiences but also lays the foundation for long-term outcomes into adulthood. Poor mental health during childhood through adolescence can lead to a range of adverse consequences, including academic underachievement, substance abuse, and social difficulties (Fletcher, 2010; Liu et al., 2011). Children with poor mental health may have adverse external effects on their community (Kauten and Barry, 2020) and on their mothers' labor supply decisions (Coley et al., 2011; Richard et al., 2014).

The importance of socio-environmental factors in children's and adolescents' mental health is well-documented. Studies in economics, sociology, psychology, and public health document that socio-environmental factors, such as social support, major and minor stressful events, and experiencing poverty, have significant impacts on mental health, encompassing self-esteem, reporting and being treated for emotional and behavioral problems (i.e., DuBois et al., 1994; Brooks-Gunn and Duncan, 1997; Gleason et al., 2016). Since many children and adolescents spend a considerable amount of time in school, in this paper, we focus on the mental health of students from late elementary school to high school.

Most education systems have a single birth date cutoff for entering school in the current academic year. The single school entry cutoff date creates variation in age among peers within a grade. As a result, how students' birth dates relate to the school entry cutoff

date could have a substantial impact on experience and achievement in school and even later in life. Studies have shown that relatively older students perform better academically, socially, behaviorally, and psychologically. For example, older students tend to have higher test scores than do their younger peers. They also demonstrate more leadership and report higher levels of confidence, risk tolerance, and grit, less ADHD diagnosis, fewer emotional symptoms, fewer peer problems, and fewer adolescent risky health behaviors (i.e., Dhuey and Lipscomb, 2008; Evans et al., 2010; Mühlenweg et al., 2012; Patalay et al., 2015; Cascio and Schanzenbach, 2016; Peña and Duckworth, 2018 Dhuey et al., 2019; Page et al., 2019; Johansen, 2021).

In this paper, we provide quasi-experimental estimates of the impact of the school entry cutoff date on students' mental health by leveraging a unique reform in South Korea that moved the school entry cutoff date from March 1 to January 1. All children in South Korea are obliged to enter elementary school if they have turned six years old by the school entry cutoff date, with a possibility of entering early or delayed. Before the reform, a substantial fraction of students born in January and February delayed entry because their parents wanted their children to have the same Korean age as their grade peers.[1] In 2008, an amendment to the Elementary and Secondary Education Act (ESEA) moved the school entry cutoff date from March 1 to January 1, effective in 2009. A few years after the reform, almost no one delayed entry.

Our paper broadens the existing empirical bases of the school starting age literature by shedding light on a less-explored context, South Korea, a Confucian-influenced East Asian country characterized by a strong age-based hierarchy. The cultural context plays an important role in understanding the effects of school starting age on mental health. Studies have shown that some aspects of culture, such as the level of individualism versus collectivism

[1]In South Korea, the traditional method of calculating age involves considering every individual to be one year old at birth and adding another year when the calendar transitions to January 1, meaning that a child born on December 31 is considered to be two years old the following day. This practice, known as the "Korean age" calculation, establishes that individuals born within the same calendar year are regarded as being the same age.

or societal equality versus hierarchy, affect how age relative to peers affects self-perceptions, decisions of cooperation, and social behaviors (Larsen et al., 2013; Bleidorn et al., 2016; Chao and Shin, 2022). All these elements might result in finding different effects of relative age on adolescent mental health in different settings (DuBois et al., 1994; Sowislo and Orth, 2013). In South Korea, small differences in biological age can have salient social impacts. For instance, Koreans often use honorifics even when addressing individuals who are one year older than themselves. Research indicates that Korean students demonstrate a heightened awareness of age differences compared to their American counterparts, particularly in how they perceive peers who are either one year older or younger (Lim and Giles, 2007). Thus, our investigation into the effects of the cutoff date reform within this distinctive cultural framework offers valuable insights into the interplay between educational policy and emotional development.

Using a difference-in-differences empirical strategy and combining it with detailed survey questionnaires on multiple dimensions of mental health, we find that the reform improves the mental health of adolescents who were born in January and February. Under the prior rule, these students would have either delayed entry or had a different Korean Age compared to their peers. We observe statistically significant effects for both girls and boys; girls exhibit 0.28 SD lower levels of inattention and 0.36 SD lower levels of aggression, and boys show 0.23 SD higher levels of self-esteem. For the group of students who would have complied with the prior age rule, the sign of the estimates is consistent with the relative age literature. The effects may be explained by these students becoming relatively older after the reform or by these students now having the same Korean age as their grade peers. For the group of students who would have delayed entry, the estimates may be explained by these students feeling less alienated and a higher sense of belonging as they no longer see themselves as the "delayed entrants."

Our paper contributes to the growing economics literature about the determinants of mental health by making use of the exogenous variation created by this policy change. Prior studies have shown that unconditional cash transfer, in utero exposure to distress, economic downturns, and introduction of social media affect children's and adolescents' mental health (Haushofer and Shapiro, 2017; Persson and Rossin-Slater, 2018; Golberstein et al., 2019; Donati et al., 2022; Braghieri et al., 2022). We focus on an element of socio-environments that all students will experience, having a different age from their classmates, rather than extreme distress such as family rupture or an event that can only happen once such as the introduction of the internet or social media. Thus, our findings have implications for a broader population.

Our paper also contributes to the literature on school starting age by using a difference-in-difference design. Many education systems have employed a single school-entry age cutoff, resulting in age variation within grade cohorts. The early literature, inspired by Angrist and Keueger (1991), commonly used birth month variation as instruments for years of schooling to study the return to education. However, concerns have been raised regarding the validity of using birth months as an instrument due to potential confounding factors such as the direct effects of the seasonality of birth or endogenous selection into certain birth months (Bound et al., 1995; Cascio and Lewis, 2006; Buckles and Hungerman, 2013). More recent literature has addressed these issues by leveraging exact birth dates and employing regression discontinuity designs (i.e., McCrary and Royer, 2011). Focusing on individuals born near the cutoff dates improves the estimates by circumventing the criticisms regarding the seasonality of births. To the best of our knowledge, the only other study in economics looking at the effects of school starting age on mental health uses the linked Danish survey and registered data to estimate the effect of kindergarten entry age on the mental health of students between ages seven and age eleven (Dee and Sievertsen, 2018). Mental health outcomes in this study are measured using self-reported scales from the Strengths and Difficulties Questionnaire developed by

English child psychiatrist Robert N Goodman in the mid-1990s, which consists of 25 items (Goodman, 1997). Using a fuzzy regression-discontinuity design based on kindergarten entry cutoff date and exact dates of birth, the authors find that a one-year delay in kindergarten entry reduces inattention/hyperactivity at age seven for girls and persists at age eleven.

The remainder of the paper is organized as follows: section 3.2 provides institutional background on education and mental health in South Korea; section 3.3 describes the data and sample used in this analysis; section 3.4 discusses empirical strategy; section 3.5 presents the results; section 3.6 concludes.

3.2. Institutional Background

3.2.1. The Traditional Korean Age-Counting Custom. In South Korea, the traditional age-counting custom ("Korean Age") considers every person one-year-old at birth and adds another year when the calendar hits January 1. This means that a child born on December 31 is one year old at birth in Korean Age and turns two the next day as the new calendar year begins. Following the Korean age calculation tradition, people born within the same calendar year are considered the same age. Even though South Korea's legal age-counting scheme follows the international age scheme after a National Law on Standardizing International Age took effect on June 28, 2023, many people in Korea are still more familiar with traditional age-counting customs.

3.2.2. School Entry Practice and Cutoff Amendment in South Korea. By law, all children in South Korea should enter elementary school in the year following the year of their sixth birthday. The academic calendar commences in March. The school entry cutoff was set to March 1 and moved to January 1 by the Elementary and Secondary Education Act (ESEA) amendment in 2008. For example, children born between March 1997 and February 1998 entered school in March 2004, while children born between January 2005 and December 2005 entered school in March 2012. That said, children born in January and

February entered school at age six with the March 1 cutoff and at age seven with the January 1 cutoff.

It is possible to enter school early or late. Early entry is rare, with around 0.3 to 0.7 percent across years at the population level before the amendment. There was an increase in early entry patterns to 2.1 percent right after the amendment because some parents wanted to send January and February-born children to school in the year of their sixth birthday, same as the March 1 cutoff. The share of early entrants decreased to 0.9 percent in 2010 and further to 0.3 percent by 2015.[2] However, delayed entry was relatively common before the amendment.[3] When March 1 served as the school entry cutoff, children born in January and February had a younger Korean age compared to their grade peers. Consequently, parents of children born in January and February increasingly intentionally delayed their children's school entry to ensure they shared the same Korean age as their classmates. Our data suggest about 30 percent to 60 percent of the January and February-born children delayed entry before the ESEA amendment, and the share of delayed entry was higher for boys by 1.0-1.2 percentage points. To address this issue, an amendment to ESEA moved the cutoff date from March 1 to January 1, effective in 2009. After the ESEA amendment, delayed entry became uncommon. About 0.8 percent of children delayed their entry in 2010, and the share decreased to less than 0.5 percent at the population level.

The amendment effectively increased the average school starting age of students born in January and February and did not change the school starting age of students born between March and December. In our data, among those who were born in January and February, the average school starting age increased from 6.47 - 6.85 before the reform to 7.01 - 7.11 after the

[2]Source: Statistical yearbook of education.

[3]To delay school entry, parents need to file an application for reasons such as child development, and the elementary school headmaster decides whether to accept the application following the deliberation of the compulsory education management committee.

reform, depending on the subsamples.[4] There is little change in the average school starting age of students born between March and December before and after the amendment for any subsample. (See the first row of Table 3.1 and Table 3.2.) The amendment also effectively reduced the share of entrants who delayed entry. If all students start school on time, the average school starting age of January and February-born students should be around 7.08. In our data, the post-reform average school starting age ranges between 7.01 and 7.11, which is close to 7.08 and suggests an overall on-time enrollment for these students. The detailed summary statistics of actual school starting ages are reported in Table 3.4.

The amendment directly affects children born in January and February 2004 and after. For the would-have-been on-time entrants who used to be the youngest within a grade without the amendment, after the amendment, these students start school a year later and become the oldest within a grade. For the would-have-been delayed entrants who are already the oldest, after the amendment, these students begin to enter school on time.

3.2.3. Adolescent Mental Health in South Korea. Mental health encompasses our psychological and social status, influencing our cognition, emotions, behaviors, and interpersonal interactions. It plays a crucial role in managing stress, fostering relationships, and guiding decision-making processes. Adolescent mental health in South Korea is a pressing concern, underscored by alarming statistics. According to the survey conducted by the Korea Center for Disease and Prevention Agency, one out of seven teenagers has contemplated suicide in 2022, reflecting a significant mental health crisis among the youth population. Children and teens grappling with mental health issues are seeking medical assistance at an alarming rate, with nearly 210,000 individuals visiting hospitals for depression or anxiety disorders from 2019 to the early half of 2022, as reported by the National Health Insurance Service. Furthermore, the suicide rate among 12 to 14-year-olds has displayed a fluctuating

[4]The school starting age is estimated using the month of birth, the year of birth, the survey year, and the survey grade using the following equation: $(SurveyYear - SurveyGrade + 1) + 3/12 - (YOB + MOB/12)$ This is because the exact date of birth was not reported in our data.

pattern over the years, starting from 1.1 per 100,000 in 2000 and reaching its peak at 5.0 per 100,000 in 2021. Similarly, for those aged 15 to 17, the suicide rate witnessed fluctuations, with a recent increase to 9.9 per 100,000 in 2020, although it slightly decreased to 9.5 per 100,000 in 2021.[5] Suicide remains a leading cause of death among teenagers in South Korea, accounting for over two in five deaths in this age group. This trend persists into early adulthood, with suicide rates remaining distressingly high among individuals in their 20s and 30s (Ha, 2023).

3.3. Data

3.3.1. Korea Children and Youth Panel Survey. Our analysis draws from two rounds of the Korea Children and Youth Panel Survey (KCYPS), namely KCYPS 2010 and KCYPS 2018, nationally representative longitudinal datasets managed by the National Youth Policy Institute of South Korea. Students from schools across 17 province-level divisions were surveyed through face-to-face interviews by trained interviewers. Parents and/or guardians were also surveyed through phone in KCYPS 2010 and face-to-face in KCYPS 2018. The survey question for students encompasses a wide range of topics including personal development (such as time use, intellectual growth, career aspirations, social and emotional development, deviant behavior, and health) as well as environmental factors (such as family dynamics, school experiences, peer relations, local community engagement, and media/cultural activities). Similarly, both KCYPS 2010 and 2018 cover background characteristics such as education, employment, and economic status in the parents' survey. Overall, the KCYPS datasets provide invaluable insights into the multifaceted experiences and trajectories of children and adolescents in South Korea, offering a holistic understanding of various facets of their lives encompassing education, health, and overall well-being.

[5]Source: Children and Youth Well-being in Korea 2022, published by Statistics Korea.

Our current analysis is based on data through the 2020 survey. We only observe cohorts who entered school in 2015 until they were in grade 6 and who entered school in 2012 until they were in grade 9.

3.3.2. Analysis Sample. Our main analytical sample comprises childhood and adolescent cohort drawn from both KCYPS 2010 and KCYPS 2018. Specifically, the childhood cohort consists of students first surveyed at grade 4 (age 10), while the early adolescence cohort comprises students first surveyed at grade 7 (age 13). Recognizing the influence of age and gender on developmental stages, we divide our data into four subsamples: child girls, child boys, early-adolescent girls, and early-adolescent boys. Each subsample is analyzed separately to account for potential differential effects across cohort-gender groups. Combining KCYPS 2010 and KCYPS 2018 panels, there are around 7000 observations per subsample. Each student-grade serves as one observation in our analysis, and cross-sectional weights are applied to ensure that sample estimates accurately reflect the population.

Two cohorts in KCYPS 2010 start school before the school entry cutoff date amendment, adhering to the March 1 cutoff rule. Two cohorts in KCYPS 2018 start school after the amendment, adhering to the January 1 cutoff rule. We exclude the cohort that entered school in 2010 in KCYPS 2010 because there is no comparable cohort in KCYPS 2018. In addition, they entered school when the reform recently took place and had different compliance patterns from other years.

3.3.3. Mental Health Outcomes. KCYPS surveyed various aspects of students' mental health in detail. We select outcomes surveyed in both KCYPS 2010 and KCYPS 2018 using the same set of questions; the outcomes are self-reported and measured on a scale. In this paper, we focus on two main types of mental health outcomes: emotional/behavioral problems and self-esteem. In particular, we focus on five emotional/behavioral problems, including both externalizing and internalizing problems: aggression, inattention (hyperactivity), social withdrawal, physical symptoms (somatic complaints), and depression.According

to Kauten and Barry (2020), "Externalizing behavior comprises any of a wide variety of generally antisocial acts (i.e., acts that violate social norms and/or are harmful to others). These acts include those that are targeted at another individual (e.g., aggression), as well as acts that may be considered victimless (e.g., substance use)." According to Hansen and Jordan (2020), "Internalizing behaviors encompass a dimension of childhood psychopathology that includes behaviors that are directed inward or are over-controlled and are associated with a number of depressive and anxiety disorders. Examples of internalizing behaviors may include fearfulness, somatic complaints, worrying, and withdrawal." As for self-esteem, we focus on global self-esteem, defined as the "level of global regard that one has for the self as a person (Harter, 1993)," in contrast to the domain-specific evaluations of self.

These outcomes are assessed with students responding to multiple questions for each construct. Using a four-point Likert scale ranging from 1 being "strongly disagree" to 4 being "strongly agree," students indicated their level of agreement with statements about various emotional problems and self-awareness constructs. The number of questions for each construct ranges from five to ten. We standardize each question, take averages, and standardize again to mean zero and unit variance.

We provide an example statement for each emotional/behavioral problem:

- Inattention: "Even after being praised or punished, I quickly become distracted."
- Aggression: "There are times when I nitpick even over small matters."
- Physical Symptoms: "I sometimes lose my appetite."
- Social Withdrawal: "I feel awkward when there are many people around."
- Depression: "I often feel low in energy."
- Self-esteem: "I feel that I have a number of good qualities."

In addition to the standardized outcomes, we also estimate a simple average of all items within each problem or self-esteem to provide descriptive statistics of the baseline mental health of children and adolescents in South Korea. We construct these simple averages by

calculating the student-level average of all items related to each problem or self-esteem. Table 3.3 present student-level descriptive statistics for constructed emotional/behavioral outcomes and self-esteem measures. Boys typically exhibit higher self-esteem than girls, and self-esteem tends to decrease during adolescence. In our sample, girls demonstrate higher levels of both externalizing and internalizing emotional/behavioral problems. Furthermore, the cohorts after the reform exhibit higher self-esteem and lower levels of emotional/behavioral problems compared to the cohorts before the reform.

One potential weakness of the survey is that not all measures are surveyed for all cohorts in all waves. For emotional/behavioral problems, data were collected from early adolescent cohorts during waves 1, 3, 5, 6, and 7 in KCYPS 2010, and during waves 1, 2, and 3 in KCYPS 2018. Similarly, data from childhood cohorts were collected during waves 3, 5, 6, and 7 in KCYPS 2010, while during waves 1, 2, and 3 in KCYPS 2018. Regarding self-esteem, early adolescent data were obtained during waves 1, 3, 5, 6, and 7 in KCYPS 2010, and during waves 1, 2, and 3 in KCYPS 2018. For students in childhood, self-esteem data were gathered during waves 2, 4, 6, and 7 in KCYPS 2010, while during waves 1, 2, and 3 in KCYPS 2018.

3.3.4. Demographics. KCYPS collected detailed demographic information about the students and their families. It is well-documented that socio-economic status (controls) affects students' emotional/behavioral problems and self-esteem (i.e., DuBois et al., 1994; Ridley et al., 2020). To control for socio-environmental differences, our analysis includes the following demographic information: whether a student is a firstborn, the educational attainment of both father and mother, and the student's home city or province at the time of the initial survey.[6] Parental education levels were categorized into four groups: less than

[6]We use the student's province of home from the first wave of each survey, minimizing the impact of selective migration.

college, 2-3 years of college, 4 years of college, and graduate degree.[7] Detailed summary statistics of these demographics for each subsample are provided in Table 3.1 and Table 3.2.

3.4. Empirical Strategy

3.4.1. Construction of the Primary Outcome Variables. In this paper, we analyze the effects of the school entry cutoff date reform on indices of six dimensions of emotional health: inattention, aggression, physical symptoms, social withdrawal, depression, and self-esteem. We construct each outcome to be an index using answers to multiple Likert-scale questions, which we will refer to as items. First, we recode all items regarding emotional/behavioral problems so that a larger value indicates a higher level of problem, and all items regarding self-esteem so that a larger value indicates a higher level of self-esteem. We check the consistency of student answers to each emotional/behavioral problem or self-esteem using Cronbach's alpha.[8] There is strong reliability, as nearly all alphas are larger than 0.8, making it reasonable for us to believe there is only one dimension and, thus, to construct one index for each emotional/behavioral problem or self-esteem.

To create each index, we first standardize each item for each cohort survey wave, compute the average of the items regarding a certain emotional/behavioral problem or self-esteem for each student, and then re-standardize the average to have a zero mean and unit variance. The initial standardization scales each item for each survey-wave-cohort, preserving as much variation as possible and ensuring uniform variation across all item-survey-wave-cohorts. This process ensures that the variation of the constructed indices is not driven by a few items with large variations or by certain survey-wave cohorts whose responses have large variations. Then, we construct the index as a simple average of the standardized items

[7] A negligible proportion, approximately 1.2 percent, of students' fathers and mothers reported education levels below high school, leading to the consolidation of the "less than high school" and "high school graduate" categories.

[8] Cronbach's alpha, developed by Lee Cronbach in 1951 measures reliability or internal consistency. $\alpha = \frac{N \times \bar{c}}{\bar{v} + (N-1) \times \bar{c}}$, where N is the number of items; $\bar{c}$ is the average covariance between item-pairs; $\bar{v}$ is the average variance. We estimate the Cronbach's alpha for each construct for each cohort in each survey wave using the stata command alpha.

because of the high consistency of answers regarding each problem or self-esteem. We re-standardize the averages to have a zero mean and unit variance for easier interpretation of regression coefficients and comparison of effect sizes with the literature. These standardized indices are used as outcomes for our regression analysis. We exclude any observations with at least one non-response item within each problem or self-esteem.[9]

3.4.2. Baseline Model. We estimate the effects of the reform on mental health using a Difference-in-Differences model. We separately estimate the effects of each survey cohort by gender, allowing for heterogeneity between gender and cohorts. Our baseline model is specified as:

$$Y_{isg} = \beta_1 Post_s + \beta_2 Treat_i + \beta_3 Post_s \times Treat_i + \theta_g + X_i\Gamma + \epsilon_{isg}$$

Here, Y_{isg} denotes the mental health outcomes of student i in survey s with currently in grade g, encompassing emotional problems (attention, aggression, physical symptoms, social withdrawal, and depression) and self-esteem. The binary variable $Post_s$ equals one if the student participated in the KCYPS 2018 survey(i.e., started school after the reform) and zero if in the KCYPS 2010 survey. Because all students in KCYPS 2010 started school before the reform and all students in KCYPS 2018 started school after the reform, the binary variable $Post_s$ also indicates if starting school after the starting subsample cutoff reform. The binary variable $Treat_i$ equals one if a student was born in January or February and zero if a student was born in March to December. The interaction term $Post_s \times Treat_i$ takes on value one if a student was born in January and February and participated in the KCYPS 2018, zero otherwise. Our main coefficient of interest is β_3, which captures the effects of being born in January or February after the reform relative to before. Because the outcomes are standardized indices, the units for the estimated effects is in standard deviations. We control for grade fixed effect (θ_g) and a set of individual demographics (X_i). ϵ_{isg} represents

[9]As a result, we exclude 2 to 5 observations, depending on the outcome or subsample.

the error term, and standard errors are clustered at the school-survey year level. This accounts for potential correlations in unobserved shocks among students attending the same school in the same year. Each observation is weighted using cross-sectional survey weights to make the sample estimates represent the population.

3.4.3. Identifying Assumptions. The effect of the cutoff date reform on mental health for January-to-February-born students relative to the March-to-December-born students can be identified using β_3 from our Difference-in-Differences design under the assumptions that: (1) the March-to-December-born students are a valid comparison group for the January-and-February-born students; (2) the month of birth effects, including selection effects and physiological effects, were unchanged because of the reform. Our settings and our balance tests suggest that it is plausible that both the above assumptions are satisfied.

First, we conduct a balance test comparing predetermined characteristics between students born in January and February and those born in March through December to check if there are systematic differences between these two groups. The characteristics include whether a student is a first-born, the student's mother's education, the student's father's education, and the home city or province of a student. These characteristics are expected not to be correlated with whether a student was born in January or February. We estimate the following regressions for each characteristic

$$X_i = \delta_0 + \delta_1 Treat_i + \epsilon_i$$

where X_i represents the characteristic and $Treat_i$ is a binary variable indicating January or February born. The results are summarized in Table 3.4, where each cell displays coefficients from separate regressions for specific samples. Overall, we find minimal differences between January-to-February-born students and March-to-December-born students, both statistically and economically, with a few exceptions. Notably, among early adolescent boy cohort, January and February births are associated with a higher proportion of fathers with

graduate degrees and a lower proportion with 2-3 year college degrees, potentially influenced by multiple tests. However, these exceptions aside, our findings suggest the predetermined characteristics are balanced between these two groups in all subsamples.

Additionally, we assume that the selection into birth months remains unchanged before and after the reform. This is plausible because our sample in both KCYPS 2010 and KCYPS 2018 were born before the reform was announced. We also assume that the direct effects of birth season remain consistent before and after the reform. It is plausible because the physiological effects of the birth season are unlikely to change drastically within a few years.

To interpret the estimated effect of the reform on January-to-February-born students relative to March-to-December-born students, it is important to underscore the high fraction of students who would have delayed entry even without the reform. These students always choose to be relatively older in their grade cohort. In our sample, 40 percent to 70 percent of January-to-February-born students voluntarily delayed school entry by a year under the March 1 cutoff. In other words, the school starting age of 40 percent to 70 percent in the treated group was not affected by the reform. The rest, 30 to 60 percent of January-to-February-born students become seven years old instead of six years old.

Therefore, the estimated treatment effect is the weighted average of two different types of treatments. The first type of treatment arises from the change in mindset of students who would start school at seven years old regardless of the cutoff. Around 40 percent to 70 percent of the students from the treatment group experienced this type of treatment, depending on the subsample. Analysis in an earlier version of Shin (2023) shows that students who delayed elementary school entry exhibit lower social-emotional skills in middle school and high school relative to students who entered school on time. One potential explanation could be that students who were held back for a year were self-conscious, leading to worse emotional and psychological well-being. After the reform, even though the school starting age was not affected, their perception of themselves changed from "delayed" to "on time," which might

have played an important role in mental health during late childhood and early adolescence. For this type of treatment effect, only the January-to-February-born students' self-perception was affected.

The second type of treatment arises from the change in relative age. As the reform moved the school entry cutoff date from March 1 to January 1, students started school one year later and became relatively older within the classroom. In our sample, around 30 percent to 60 percent of the January-to-February-born students experienced this type of treatment, depending on the subsample. As students became relatively older, the students who were born between March and December became relatively younger after the reform. As the comparison group was also affected by the treatment, the interpretation of the treatment effects of relative age should be approached with caution. We present a back-of-the-envelope calculation to conceptualize the interpretation. We simplified the calculation by assuming the effects only come from the rank of birth month and that the effect of the ordinal rank of the birth month within the classroom on mental health outcomes is linear. Students' birth months rank from 1st to 12th, with 1st being the birth month of the oldest student. As a result of this reform, the birth month rank of January- and February-born students whose age increases to seven years old moved up by 10: the January-born students moved from the 11th to the 1st, and the February-born students moved from the 12th to the 2nd. The rank of the students who were born in March to December moved down by 2: for example, the March-born students moved from 1st to 3rd. The estimated effects of our specification capture the combined effects of the treatment group becoming relatively older and the comparison group becoming relatively younger. In other words, the effect of a one-year change in relative age. To interpret the effects of the treatment on the January-to-February-born students, the estimated coefficient should be scaled down by 5/6.

3.5. Results

We summarize the coefficients from our baseline specification on mental health indices in Table 3.5. Each panel consists of a different subsample. Within each panel, each column represents the coefficients of a different outcome, with the outcome indicated in the column name. The first row of each panel shows the estimates of β_1, the second row shows the estimates of β_2, and the third row shows the estimates of β_3. As the outcomes are standardized, all the estimated coefficients should be interpreted as multiples of standard deviations (SD). For our key coefficient of interest, β_3, our estimates suggest that early-adolescent girls born in January and February indicate substantially lower levels of external emotional/behavioral problems after the reform, and early-adolescent boys born in January and February indicate substantially higher levels of self-esteem after the reform. There are no statistically significant differences in any outcomes for boys and girls in their childhood. We also find evidence of trends in mental health over time and supporting evidence of the importance of separately analyzing mental health by gender and developmental stages.

Panel A of Table 3.5 summarizes the estimates of mental health indicators using the early-adolescent girl sample. We find that early-adolescent girls born in January and February who started school after the reform exhibit a 0.28 SD lower level of inattention and a 0.36 SD lower level of aggression relative to before. Additionally, our estimates suggest that compared to the early-adolescent girls who participated in KCYPS 2010, those in KCYPS 2018 exhibited higher levels of aggression, lower levels of self-esteem, and suggestively higher levels of other emotional/behavioral problems. Before the reform, early-adolescent girls born in January and February exhibited higher levels of aggression and suggestively higher levels of inattention relative to girls born between March and December.

Panel B of Table 3.5 summarizes the estimates using the early-adolescent boy sample. We find that early-adolescent boys born in January and February, who started school after the reform, exhibit a 0.22 SD higher level of self-esteem relative to before. We find no

statistically significant differences in any emotional/behavioral problems. Our estimates also indicate that compared to early-adolescent boys who participated in KCYPS 2010, those in KCYPS 2018 showed an overall higher level of depression and potentially higher levels of other emotional/behavioral problems. Before the reform, early-adolescent boys born in January and February exhibited higher levels of social withdrawal and potentially lower levels of self-esteem relative to boys born between March and December. The observed differences in mental health effects between boys and girls in early adolescence underscore the importance of conducting separate analyses by gender to comprehend adolescent mental health fully. Together with our findings in Panel A, our estimates suggest that the reform has substantial effects on mental health for both boys and girls in early adolescence, with the effects on girls primarily manifesting through externalizing emotional/behavioral problems and the effects on boys primarily through self-esteem.

Why do we see these differences? First, regarding students born in January and February who started school one year later due to the reform, they may exhibit improved mental health because they are relatively the oldest in their cohort, as suggested by the relative age literature on mental health and ADHD diagnosis (Evans et al., 2010; Dee and Sievertsen, 2018). One potential mechanism through which relative age affects mental health could be that January and February-born students are less likely to experience delayed physical development compared to their peers through puberty after the reform. Second, considering the high fraction of students who delayed entry before the reform, we also hypothesize that parents' decisions to 'delay entry' may have a negative effect on students' mental health during early adolescence. This hypothesis suggests that despite red-shirting being intended to provide boys with advantages in academic and social development, delayed entry into school may lead to feelings of inadequacy or low self-worth, potentially adversely affecting their self-esteem. Further analysis is needed to better understand the effect of intentionally delayed entry on students' mental health.

Panel C of Table 3.5 summarises the estimates of mental health indicators using the child girl sample. We find that young girls who were born in January and February and started school after the reform exhibit no statistically significant differences in emotional/behavioral problems or self-esteem relative to before the reform. Additionally, our estimates indicate that compared to childhood girls who participated in KCYPS 2010, those in KCYPS 2018 exhibit lower levels of emotional/behavioral problems, such as physical symptoms, social withdrawal, and depression. Before the reform, childhood girls born in January and February exhibit lower levels of self-esteem and higher levels of social withdrawal relative to girls born in March and December. The observed differences in estimated effects between young girls and early adolescence underscore the importance of conducting separate analyses by developmental stages to understand students' mental health.

Panel D of Table 3.5 summarises the estimates of mental health indicators using the child boy sample. We find that young boys who were born in January and February and started school after the reform do not exhibit statistically significant differences in levels of emotional/behavioral problems or self-esteem relative to those starting school before the reform. Additionally, our estimates indicate that compared to child boys who participated in KCYPS 2010, those in KCYPS 2018 exhibit higher levels of emotional/behavioral problems, such as physical symptoms, social withdrawal, and depression. Before the reform, child boys born in January and February exhibit little difference in emotional/behavioral problems and self-esteem relative to boys born in March and December. From Panel C and D, the overall small difference we found in both young boys and young girls is consistent with the findings of the literature on diminishing effects on non-cognitive skills from early childhood to late childhood (Datar and Gottfried, 2015; Lubotsky and Kaestner, 2016).

Compared to the existing literature on mental health, the magnitude of our estimates is consistent with prior findings, suggesting substantial impacts of the reform. Relative to Dee and Sievertsen (2018), which employs a fuzzy RD design with Danish data and demonstrates

that one-year later entry reduced inattention/hyperactivity by 0.7 SD at age seven and 0.68 SD at age eleven, we observed no effects around ages 10 to 13 but found effects half the size for students around ages 13 to 16.

3.6. Conclusion

In this paper, we estimate the impact of a South Korean school starting age cutoff reform, which moved the cutoff from March 1 to January 1, on students' mental health outcomes, encompassing both internalizing and externalizing emotional/behavioral problems and self-esteem. We draw on comprehensive data on measures including inattention, aggression, physical symptoms, social withdrawal, depression, and self-esteem surveyed among students at various developmental stages in the Korea Children and Youth Panel Survey (KCYPS). Using a difference-in-differences model, our analysis suggests that the reform substantially improved the mental health of students born in January and February.

In particular, we find that early-adolescent girls born in January and February who started school after the reform report lower levels of inattention and aggression compared to before. Early-adolescent boys born in January and February after the reform report higher levels of self-esteem relative to before. Additionally, our analysis uncover nuanced effects across different age groups and genders, underscoring the importance of considering development stage and gender in policy formulation. We hypothesize that the relative age within the classroom and the mindset of delayed entrants might explain the observed effects.

Looking ahead, our findings have significant implications for policymakers and educators aiming to enhance mental health outcomes among students. By acknowledging heterogeneous impacts of educational reform on mental health across gender and developmental stages, education policymakers and practitioners can design interventions tailored to address specific needs and vulnerabilities across diverse demographic groups. Future research into students' mental health and the underlying mechanisms driving these effects is crucial for informed education policy-making and the promotion of student well-being.

3.7. Figure

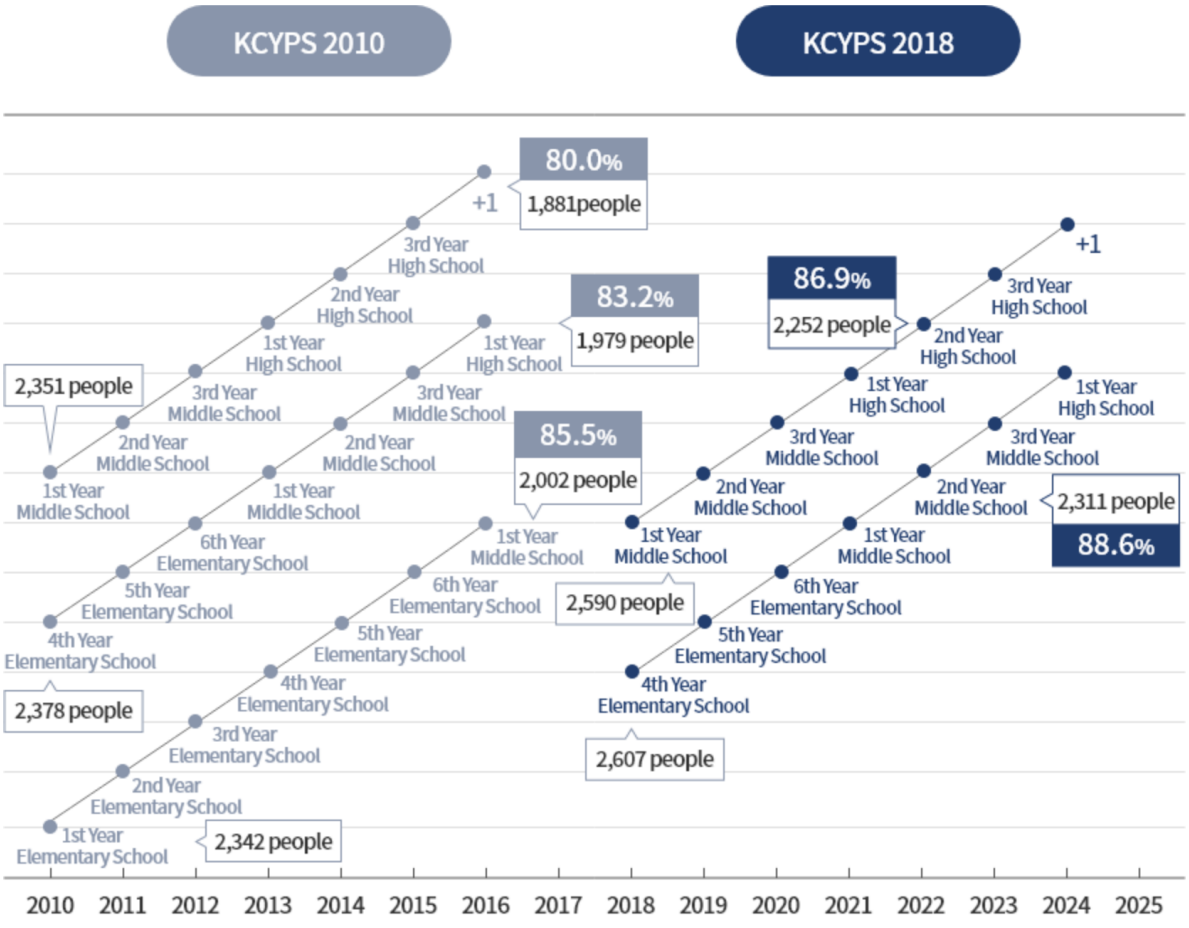


FIGURE 3.1. Korea Children and Youth Panel Survey Cohorts

3.8. Tables

TABLE 3.1. Descriptive Statistics of Predetermined Characteristics: Early Adolescence Cohorts

	Girls				Boys			
	Before reform		After reform		Before reform		After reform	
	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec
School starting age	6.466	6.620	7.010	6.644	6.699	6.647	7.087	6.643
	(0.463)	(0.290)	(0.316)	(0.240)	(0.487)	(0.288)	(0.224)	(0.250)
Firstborn	0.450	0.484	0.460	0.516	0.513	0.491	0.421	0.479
	(0.499)	(0.500)	(0.500)	(0.500)	(0.501)	(0.500)	(0.495)	(0.500)
Mother's education								
Less than college	0.427	0.488	0.293	0.270	0.503	0.475	0.347	0.290
	(0.496)	(0.500)	(0.456)	(0.444)	(0.501)	(0.500)	(0.477)	(0.454)
2-3 year college graduate	0.147	0.113	0.274	0.293	0.061	0.132	0.276	0.256
	(0.355)	(0.317)	(0.447)	(0.456)	(0.240)	(0.338)	(0.448)	(0.436)
4 year college graduate	0.390	0.366	0.363	0.386	0.372	0.364	0.343	0.372
	(0.489)	(0.482)	(0.482)	(0.487)	(0.485)	(0.482)	(0.476)	(0.484)
Graduate degree	0.036	0.033	0.071	0.050	0.064	0.029	0.033	0.083
	(0.187)	(0.178)	(0.257)	(0.219)	(0.246)	(0.169)	(0.180)	(0.275)
Father's education								
Less than college	0.341	0.369	0.259	0.259	0.366	0.367	0.322	0.256
	(0.475)	(0.483)	(0.439)	(0.438)	(0.483)	(0.482)	(0.468)	(0.437)
2-3 year college graduate	0.105	0.081	0.163	0.205	0.046	0.086	0.125	0.189
	(0.307)	(0.273)	(0.370)	(0.404)	(0.211)	(0.280)	(0.331)	(0.392)
4 year college graduate	0.489	0.487	0.459	0.422	0.451	0.474	0.451	0.442
	(0.501)	(0.500)	(0.500)	(0.494)	(0.499)	(0.500)	(0.499)	(0.497)
Graduate degree	0.065	0.064	0.119	0.113	0.137	0.073	0.103	0.112
	(0.247)	(0.244)	(0.325)	(0.317)	(0.345)	(0.260)	(0.304)	(0.316)
City/province of home								
Seoul	0.148	0.178	0.155	0.163	0.203	0.168	0.239	0.147
	(0.355)	(0.383)	(0.363)	(0.370)	(0.403)	(0.374)	(0.427)	(0.354)
Busan	0.106	0.053	0.031	0.064	0.043	0.068	0.036	0.061
	(0.309)	(0.224)	(0.173)	(0.244)	(0.203)	(0.252)	(0.186)	(0.239)
Daegu	0.052	0.053	0.056	0.047	0.047	0.056	0.038	0.051
	(0.222)	(0.224)	(0.230)	(0.211)	(0.212)	(0.230)	(0.192)	(0.220)
Incheon	0.040	0.061	0.036	0.062	0.055	0.056	0.058	0.056
	(0.196)	(0.239)	(0.186)	(0.241)	(0.228)	(0.230)	(0.235)	(0.230)
Gwangju	0.028	0.037	0.042	0.028	0.011	0.040	0.025	0.036
	(0.167)	(0.189)	(0.200)	(0.166)	(0.106)	(0.196)	(0.158)	(0.185)
Daejeon	0.033	0.033	0.036	0.032	0.030	0.031	0.032	0.033
	(0.180)	(0.179)	(0.186)	(0.176)	(0.171)	(0.173)	(0.176)	(0.178)
Ulsan	0.034	0.024	0.029	0.022	0.023	0.027	0.014	0.026
	(0.182)	(0.153)	(0.169)	(0.147)	(0.149)	(0.163)	(0.117)	(0.160)
Sejong	0.000	0.000	0.009	0.007	0.000	0.000	0.011	0.006
	(0.000)	(0.000)	(0.096)	(0.082)	(0.000)	(0.000)	(0.106)	(0.076)
Gyeonggi-do	0.260	0.246	0.244	0.281	0.272	0.240	0.267	0.271
	(0.440)	(0.431)	(0.430)	(0.450)	(0.446)	(0.428)	(0.443)	(0.445)
Gangwon-do	0.033	0.030	0.070	0.021	0.020	0.031	0.040	0.028
	(0.178)	(0.171)	(0.255)	(0.144)	(0.140)	(0.174)	(0.196)	(0.165)
Chungcheongbuk-do	0.034	0.032	0.023	0.033	0.026	0.035	0.029	0.032
	(0.183)	(0.175)	(0.149)	(0.178)	(0.161)	(0.184)	(0.167)	(0.177)
Chungcheongnam-do	0.029	0.043	0.056	0.040	0.051	0.038	0.034	0.043
	(0.167)	(0.203)	(0.230)	(0.195)	(0.221)	(0.191)	(0.182)	(0.202)
Jeollabuk-do	0.037	0.039	0.045	0.041	0.055	0.035	0.041	0.038
	(0.189)	(0.193)	(0.207)	(0.198)	(0.228)	(0.185)	(0.198)	(0.190)
Jeollanam-do	0.043	0.040	0.046	0.033	0.035	0.039	0.017	0.039
	(0.203)	(0.195)	(0.210)	(0.178)	(0.184)	(0.194)	(0.130)	(0.194)
Gyeongsangbuk-do	0.040	0.051	0.040	0.047	0.051	0.053	0.028	0.052
	(0.197)	(0.220)	(0.197)	(0.212)	(0.221)	(0.224)	(0.165)	(0.223)
Gyeongsangnam-do	0.078	0.066	0.061	0.068	0.061	0.069	0.071	0.069
	(0.269)	(0.249)	(0.239)	(0.252)	(0.241)	(0.254)	(0.258)	(0.253)
Jeju-do	0.005	0.014	0.023	0.012	0.017	0.012	0.021	0.013
	(0.073)	(0.120)	(0.151)	(0.109)	(0.130)	(0.110)	(0.143)	(0.114)
Observations	209	966	193	992	202	974	228	1177

Note: Standard deviation in parentheses. Early adolescence cohorts are defined as the cohorts who first participated in the survey at grade seven. Before the reform estimates are from students who participated in the KCYPS 2010. After the reform estimates are from students who participated in KCYPS 2018.

TABLE 3.2. Descriptive Statistics of Predetermined Characteristics: Childhood Cohorts

	Girls				Boys			
	Before reform		After reform		Before reform		After reform	
	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec
School starting age	6.796	6.624	7.105	6.637	6.848	6.630	7.113	6.629
	(0.472)	(0.245)	(0.139)	(0.252)	(0.447)	(0.252)	(0.139)	(0.248)
Firstborn	0.467	0.476	0.460	0.425	0.432	0.445	0.475	0.490
	(0.500)	(0.500)	(0.500)	(0.495)	(0.496)	(0.497)	(0.501)	(0.500)
Mother's education								
Less than college	0.398	0.407	0.220	0.244	0.439	0.446	0.251	0.214
	(0.491)	(0.491)	(0.415)	(0.430)	(0.498)	(0.497)	(0.435)	(0.410)
2-3 year college graduate	0.131	0.143	0.306	0.284	0.160	0.152	0.276	0.285
	(0.339)	(0.350)	(0.462)	(0.451)	(0.367)	(0.359)	(0.448)	(0.451)
4 year college graduate	0.419	0.406	0.375	0.390	0.380	0.366	0.410	0.409
	(0.495)	(0.491)	(0.485)	(0.488)	(0.486)	(0.482)	(0.493)	(0.492)
Graduate degree	0.052	0.045	0.099	0.082	0.021	0.037	0.063	0.092
	(0.222)	(0.207)	(0.300)	(0.275)	(0.144)	(0.188)	(0.243)	(0.290)
Father's education								
Less than college	0.309	0.325	0.248	0.222	0.388	0.392	0.218	0.181
	(0.463)	(0.469)	(0.433)	(0.416)	(0.488)	(0.488)	(0.414)	(0.385)
2-3 year college graduate	0.123	0.100	0.152	0.215	0.124	0.088	0.198	0.218
	(0.329)	(0.300)	(0.360)	(0.411)	(0.331)	(0.283)	(0.399)	(0.413)
4 year college graduate	0.493	0.492	0.464	0.429	0.414	0.446	0.450	0.489
	(0.501)	(0.500)	(0.500)	(0.495)	(0.494)	(0.497)	(0.498)	(0.500)
Graduate degree	0.075	0.082	0.136	0.133	0.074	0.075	0.134	0.112
	(0.263)	(0.275)	(0.344)	(0.340)	(0.262)	(0.263)	(0.342)	(0.316)
City/province of home								
Seoul	0.160	0.175	0.197	0.150	0.184	0.168	0.141	0.163
	(0.367)	(0.380)	(0.399)	(0.357)	(0.389)	(0.374)	(0.349)	(0.370)
Busan	0.096	0.052	0.048	0.058	0.079	0.057	0.047	0.059
	(0.295)	(0.222)	(0.213)	(0.234)	(0.270)	(0.231)	(0.212)	(0.235)
Daegu	0.065	0.047	0.048	0.047	0.062	0.049	0.046	0.047
	(0.248)	(0.211)	(0.214)	(0.212)	(0.243)	(0.216)	(0.209)	(0.212)
Incheon	0.052	0.056	0.027	0.067	0.077	0.050	0.082	0.052
	(0.222)	(0.231)	(0.161)	(0.250)	(0.267)	(0.218)	(0.274)	(0.223)
Gwangju	0.043	0.035	0.048	0.029	0.035	0.035	0.022	0.038
	(0.204)	(0.184)	(0.214)	(0.169)	(0.185)	(0.183)	(0.146)	(0.190)
Daejeon	0.043	0.031	0.031	0.032	0.023	0.035	0.036	0.030
	(0.203)	(0.174)	(0.175)	(0.176)	(0.151)	(0.184)	(0.186)	(0.171)
Ulsan	0.023	0.025	0.005	0.029	0.024	0.025	0.021	0.026
	(0.151)	(0.156)	(0.072)	(0.168)	(0.154)	(0.157)	(0.143)	(0.159)
Sejong	0.000	0.000	0.014	0.007	0.000	0.000	0.010	0.006
	(0.000)	(0.000)	(0.118)	(0.083)	(0.000)	(0.000)	(0.100)	(0.079)
Gyeonggi-do	0.205	0.268	0.258	0.282	0.196	0.268	0.297	0.269
	(0.405)	(0.443)	(0.438)	(0.450)	(0.398)	(0.443)	(0.458)	(0.444)
Gangwon-do	0.033	0.029	0.040	0.025	0.036	0.029	0.034	0.028
	(0.180)	(0.168)	(0.196)	(0.156)	(0.186)	(0.168)	(0.183)	(0.164)
Chungcheongbuk-do	0.036	0.031	0.048	0.027	0.039	0.030	0.009	0.038
	(0.187)	(0.172)	(0.214)	(0.161)	(0.195)	(0.171)	(0.093)	(0.191)
Chungcheongnam-do	0.059	0.039	0.041	0.044	0.046	0.041	0.055	0.040
	(0.236)	(0.193)	(0.198)	(0.204)	(0.210)	(0.199)	(0.228)	(0.196)
Jeollabuk-do	0.028	0.041	0.040	0.036	0.031	0.040	0.031	0.038
	(0.166)	(0.198)	(0.196)	(0.186)	(0.172)	(0.196)	(0.174)	(0.191)
Jeollanam-do	0.051	0.034	0.038	0.034	0.047	0.036	0.050	0.030
	(0.221)	(0.182)	(0.192)	(0.181)	(0.213)	(0.186)	(0.218)	(0.172)
Gyeongsangbuk-do	0.049	0.050	0.052	0.045	0.057	0.050	0.054	0.046
	(0.216)	(0.217)	(0.222)	(0.208)	(0.231)	(0.218)	(0.227)	(0.209)
Gyeongsangnam-do	0.037	0.074	0.050	0.075	0.052	0.074	0.057	0.074
	(0.190)	(0.262)	(0.218)	(0.263)	(0.222)	(0.262)	(0.232)	(0.262)
Jeju-do	0.019	0.013	0.016	0.014	0.011	0.014	0.009	0.016
	(0.136)	(0.113)	(0.126)	(0.119)	(0.105)	(0.116)	(0.094)	(0.125)
Observations	196	937	251	1043	219	1026	279	1034

Standard deviation in parentheses.

Note: Standard deviation in parentheses. Childhood cohorts are defined as the cohorts who first participated in the survey at grade four. Before the reform estimates are from students who participated in the KCYPS 2010. After the reform estimates are from students who participated in KCYPS 2018.

TABLE 3.3. Summary Statistics of Baseline Outcomes

	Entire sample				Girls				Boys			
	Before reform		After reform		Before reform		After reform		Before reform		After reform	
	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec	Jan - Feb	Mar - Dec
Panel A. Early Adolescence												
Inattention	2.388	2.392	2.080	2.191	2.456	2.370	2.027	2.148	2.323	2.413	2.135	2.230
	(0.586)	(0.525)	(0.556)	(0.569)	(0.559)	(0.508)	(0.555)	(0.579)	(0.605)	(0.539)	(0.553)	(0.557)
Aggression	2.165	2.129	1.851	1.942	2.289	2.149	1.824	1.977	2.048	2.111	1.879	1.911
	(0.596)	(0.582)	(0.578)	(0.601)	(0.578)	(0.568)	(0.572)	(0.621)	(0.590)	(0.595)	(0.584)	(0.581)
Physical symptoms	2.059	2.019	1.869	1.885	2.135	2.078	1.869	1.927	1.986	1.964	1.869	1.846
	(0.641)	(0.633)	(0.608)	(0.611)	(0.653)	(0.631)	(0.595)	(0.630)	(0.623)	(0.631)	(0.623)	(0.591)
Social withdrawal	2.254	2.228	2.132	2.138	2.288	2.243	2.145	2.161	2.221	2.214	2.118	2.117
	(0.711)	(0.711)	(0.777)	(0.693)	(0.706)	(0.699)	(0.819)	(0.694)	(0.715)	(0.721)	(0.733)	(0.691)
Depression	1.983	1.921	1.729	1.803	2.110	2.006	1.721	1.881	1.862	1.842	1.738	1.732
	(0.644)	(0.614)	(0.617)	(0.608)	(0.671)	(0.598)	(0.600)	(0.625)	(0.594)	(0.618)	(0.635)	(0.583)
Self-esteem	2.771	2.798	3.023	2.985	2.732	2.756	2.909	2.891	2.809	2.838	3.138	3.071
	(0.494)	(0.517)	(0.520)	(0.493)	(0.499)	(0.507)	(0.565)	(0.509)	(0.488)	(0.524)	(0.442)	(0.461)
Observations	396	1884	397	2041	198	930	181	939	198	954	216	1102
Panel B. Childhood												
Inattention	2.342	2.281	2.163	2.176	2.239	2.216	2.082	2.124	2.427	2.341	2.229	2.228
	(0.642)	(0.590)	(0.564)	(0.590)	(0.585)	(0.531)	(0.581)	(0.591)	(0.675)	(0.635)	(0.541)	(0.586)
Aggression	2.119	2.015	1.895	1.928	2.089	2.006	1.925	1.917	2.142	2.024	1.870	1.939
	(0.610)	(0.616)	(0.605)	(0.632)	(0.620)	(0.592)	(0.618)	(0.645)	(0.602)	(0.638)	(0.593)	(0.619)
Physical symptoms	1.922	1.878	1.811	1.809	1.980	1.946	1.840	1.816	1.874	1.815	1.787	1.803
	(0.690)	(0.624)	(0.616)	(0.639)	(0.721)	(0.620)	(0.619)	(0.641)	(0.662)	(0.621)	(0.614)	(0.636)
Social withdrawal	2.090	2.031	2.153	2.167	2.228	2.087	2.203	2.212	1.977	1.980	2.111	2.122
	(0.790)	(0.750)	(0.699)	(0.745)	(0.836)	(0.724)	(0.680)	(0.759)	(0.733)	(0.769)	(0.712)	(0.730)
Depression	1.753	1.696	1.741	1.755	1.826	1.788	1.823	1.790	1.694	1.610	1.672	1.721
	(0.663)	(0.592)	(0.610)	(0.617)	(0.670)	(0.605)	(0.642)	(0.632)	(0.652)	(0.567)	(0.575)	(0.600)
Self-esteem	2.987	3.022	3.090	3.079	3.017	3.037	3.067	3.082	2.962	3.008	3.110	3.077
	(0.463)	(0.461)	(0.462)	(0.470)	(0.484)	(0.462)	(0.487)	(0.493)	(0.444)	(0.460)	(0.439)	(0.448)
Observations	382	1837	486	1925	176	874	229	971	206	963	257	954

Note: Standard deviation in parentheses. The baseline is defined as a first overlapping grade (wave) across the 2010 panel and the 2018 panel. The first overlapping grade is grade 8 for inattention, aggression, physical symptoms, social withdrawal, and depression in the middle adolescent cohort; grade 7 for self-esteem in the middle adolescent cohort; grade 6 for inattention, aggression, physical symptoms, social withdrawal in the early adolescent cohort; and grade 5 for self-esteem in the early adolescent cohort. Data Source: Korea Children and Youth Panel Survey 2010 and 2018.

TABLE 3.4. Balance Test on Predetermined Characteristics

	Early adolescence			Childhood		
	Entire sample	Girls	Boys	Entire sample	Girls	Boys
Firstborn	-0.022	-0.042	-0.002	-0.002	0.007	-0.011
	(0.019)	(0.034)	(0.025)	(0.022)	(0.029)	(0.029)
Mother's education						
Less than college	0.011	-0.023	0.046	-0.012	-0.023	-0.002
	(0.021)	(0.033)	(0.032)	(0.018)	(0.030)	(0.025)
2-3 year college graduate	-0.015	0.008	-0.040*	0.011	0.011	0.010
	(0.017)	(0.024)	(0.023)	(0.015)	(0.024)	(0.022)
4 year college graduate	-0.001	0.005	-0.007	0.005	-0.001	0.010
	(0.022)	(0.031)	(0.031)	(0.022)	(0.032)	(0.028)
Graduate degree	0.005	0.009	0.001	-0.004	0.013	-0.018
	(0.013)	(0.014)	(0.019)	(0.010)	(0.017)	(0.012)
Father's education						
Less than college	0.007	-0.014	0.028	0.001	-0.000	0.002
	(0.024)	(0.030)	(0.034)	(0.018)	(0.029)	(0.026)
2-3 year college graduate	-0.029**	-0.005	-0.052***	0.003	-0.014	0.018
	(0.013)	(0.021)	(0.016)	(0.014)	(0.021)	(0.019)
4 year college graduate	0.003	0.017	-0.010	-0.011	0.014	-0.033
	(0.022)	(0.031)	(0.034)	(0.020)	(0.032)	(0.025)
Graduate degree	0.018	0.002	0.034*	0.007	-0.000	0.013
	(0.013)	(0.018)	(0.019)	(0.012)	(0.021)	(0.016)

Note: Standard errors in parentheses. Statistical significance * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the first-wave school level. Weighted using the first-wave cross-sectional weights. To check the balance of the city/province of home, we use wave 1 of KCYPS 2010 and KCYPS 2018. Data Source: Korea Children and Youth Panel Survey 2010 and 2018.

TABLE 3.5. Differential MOB Effects on Emotional Well-beings - Baseline Model

	(1) inattention	(2) aggression	(3) physical symptoms	(4) social withdrawal	(5) depression	(6) self-esteem
Panel A. Early Adolescence, Girls						
Post	0.0377	0.125**	0.0741	0.0330	0.0295	-0.132**
	(0.63)	(2.91)	(1.80)	(0.68)	(0.68)	(-2.98)
Treat	0.126	0.193**	0.0421	-0.0258	0.0418	0.0337
	(1.78)	(4.08)	(0.86)	(-0.45)	(0.55)	(0.61)
Post × Treat	-0.277**	-0.356***	-0.0517	0.0311	-0.115	0.0638
	(-3.24)	(-4.68)	(-0.55)	(0.28)	(-0.81)	(0.92)
Observations	6777	6777	6781	6779	6776	6812
Panel B. Early Adolescence, Boys						
Post	0.0648	0.0594	0.0669	0.0424	0.106*	0.0163
	(1.34)	(1.67)	(1.89)	(1.40)	(2.49)	(0.46)
Treat	-0.0442	0.0129	-0.0136	0.0965**	0.0599	-0.0771
	(-0.58)	(0.25)	(-0.26)	(2.82)	(1.47)	(-1.25)
Post × Treat	-0.0502	-0.0209	0.00355	-0.0363	-0.0585	0.226**
	(-0.60)	(-0.21)	(0.05)	(-0.58)	(-1.01)	(2.83)
Observations	7208	7214	7213	7209	7212	7222
Panel C. Childhood, Girls						
Post	-0.0242	-0.0371	-0.0997***	-0.0473**	-0.130***	-0.0415
	(-1.16)	(-1.65)	(-4.61)	(-2.58)	(-6.30)	(-1.84)
Treat	-0.0120	-0.00752	-0.0931	0.123**	-0.0209	-0.0747**
	(-0.27)	(-0.10)	(-1.55)	(2.76)	(-0.58)	(-2.85)
Post × Treat	-0.0948	0.0388	0.101	-0.0927	0.0654	0.0720
	(-1.49)	(0.43)	(1.52)	(-1.36)	(1.41)	(1.30)
Observations	6701	6701	6700	6700	6699	6742
Panel D. Childhood, Boys						
Post	0.0242	-0.00304	0.0722**	0.0580*	0.0722**	-0.0113
	(1.04)	(-0.11)	(3.38)	(2.52)	(2.79)	(-0.36)
Treat	0.0192	0.0263	-0.0288	-0.0709	-0.0201	-0.0141
	(0.35)	(0.46)	(-0.47)	(-1.35)	(-0.30)	(-0.23)
Post × Treat	-0.0899	-0.119	-0.0345	-0.0352	-0.0841	0.117
	(-1.40)	(-1.58)	(-0.47)	(-0.50)	(-0.99)	(1.61)
Observations	7036	7037	7037	7035	7037	7061

Note: t statistics in parentheses. Statistical significance * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at the school-grade level. Weighted using panel wave-specific cross-sectional weights. Data Source: Korea Children and Youth Panel Survey 2010 and 2018.

TABLE 3.6. Differntial MOB Effects on Emotional Well-Beings: Baseline Model with Pooled Sample

	Inattention		Aggression		Physical symptoms		Social withdraw		Depression		Self-esteem	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Girls												
Post	-0.009	-0.004	0.003	0.030	-0.040	-0.003	0.012	0.030	-0.062**	-0.042	0.026	0.013
	(0.026)	(0.026)	(0.021)	(0.021)	(0.022)	(0.025)	(0.019)	(0.027)	(0.022)	(0.026)	(0.031)	(0.035)
Treat	0.031	0.066	0.067	0.104**	-0.036	-0.018	0.022	0.042	-0.009	0.012	0.015	-0.011
	(0.053)	(0.053)	(0.037)	(0.041)	(0.025)	(0.024)	(0.043)	(0.046)	(0.040)	(0.043)	(0.030)	(0.033)
Post × Treat	-0.143**	-0.158**	-0.124**	-0.139**	0.062	0.046	-0.026	-0.031	-0.004	-0.014	-0.013	-0.008
	(0.059)	(0.057)	(0.039)	(0.042)	(0.040)	(0.044)	(0.062)	(0.064)	(0.053)	(0.056)	(0.035)	(0.038)
Age at survey	0.000	-0.002	0.007	0.002	0.010	0.008	0.000	-0.001	0.005	0.005	-0.002	-0.001
	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)
Observations	19,738	17,002	19,739	17,002	19,741	17,005	19,740	17,003	19,735	16,999	19,819	17,078
Panel B: Boys												
Post	0.023	0.070**	0.022	0.043**	0.055**	0.062**	0.037*	0.069***	0.100***	0.106***	-0.023	-0.061**
	(0.023)	(0.027)	(0.017)	(0.016)	(0.020)	(0.019)	(0.019)	(0.018)	(0.022)	(0.019)	(0.027)	(0.024)
Treat	-0.032	-0.009	0.023	0.025	-0.026	-0.021	0.007	0.009	0.022	0.012	-0.028	-0.042
	(0.040)	(0.041)	(0.036)	(0.034)	(0.046)	(0.041)	(0.035)	(0.033)	(0.034)	(0.033)	(0.039)	(0.046)
Post × Treat	-0.025	-0.023	-0.060	-0.029	-0.011	0.019	-0.031	-0.017	-0.075	-0.053	0.078	0.116*
	(0.057)	(0.062)	(0.048)	(0.055)	(0.051)	(0.050)	(0.047)	(0.052)	(0.047)	(0.053)	(0.046)	(0.052)
Age at survey	-0.013*	-0.017**	-0.013**	-0.015***	-0.007	-0.008	0.007	0.006	0.004	0.003	-0.001	-0.003
	(0.006)	(0.007)	(0.005)	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)	(0.004)	(0.008)	(0.007)
Observations	21,250	17,971	21,257	17,978	21,256	17,977	21,250	17,971	21,254	17,976	21,294	18,010
Individual Controls		Yes		Yes		Yes		Yes		Yes		Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX A

Data Appendix: KCYPS 2010 and KCYPS 2018

The survey employs two-stage stratified cluster sampling, ensuring a nationally representative sample. The population was first stratified into geographical divisions such as provinces and metropolitan cities. The selected schools were further stratified into sub-province levels, based on the size of cities within the geographical divisions.[1] In each stratum, schools were randomly selected in proportion to the number of classes in the stratum. Each school selected has at least 2 classes and at least 50 students. If a selected school did not meet the criteria, it was replaced with a nearby school that met both criteria. The final survey sample includes observations only when both students and parents are surveyed. While the reference year for the student population differs between the two panels, the strata are comparable. Each observation wave has a cross-sectional weight, multiplying the probability of sampling weight, the non-response weight, and the post-stratification weight. The cross-sectional weighted sample represents the population of students in the survey year.

The survey was conducted every year between September and November in KCYPS 2010 and between August and November in KCYPS 2018. Figure 3.1 illustrate the survey year, survey grade, sample size, and retention rates of all survey cohorts in both KCYPS 2010 and 2018. In the KCYPS 2010 panel, the survey tracked cohorts that entered school in 2004, 2007, and 2010, from 2010 to 2016. The cohort who entered school in 2004 were in the first year of middle school (grade 7) in the first survey wave. They were surveyed between grade 7 (ages 12-13) and one year after high school graduation (ages 18-19). The cohort who entered school in 2007 were in the fourth year of elementary school in the first survey wave and

[1]There were 27 strata in KCYPS 2010 and 28 strata in KCYPS 2018. The number of strata increased in KCYPS 2018 due to the introduction of Sejong-Si in 2012.

surveyed between grade 4 (ages 9-10) and grade 10 (ages 15-16). The cohort who entered school in 2010 were in the first year of elementary school (grade 1) in the first survey wave and surveyed until grade 7. The KCYPS 2018 panel followed cohorts that began school in 2012 and 2015, tracking them from 2018 to 2024. The cohort who entered school in 2012 were in the first year of middle school (grade 7) in the first survey wave surveyed between grade 7 and one year after high school graduation. The cohort who entered school in 2015 were in the fourth year of elementary school in the first survey wave and surveyed between grade 4 and grade 10. The retention rates for the last wave (wave 7) of KCYPS 2010 are between 80 percent to 85.5 percent, and the retention rates for wave 5 of KCYPS 2018 are 86.9 percent to 88.6 percent.

The number of questions for each construct ranges from five to ten, with seven questions for inattention, six questions for aggression, eight questions for physical symptoms, five questions for social withdrawal, ten questions for depression, and ten questions for self-esteem.

Bibliography

Akee, Randall, William Copeland, E. Jane Costello, and Emilia Simeonova, "How does household income affect child personality traits and behaviors?," *American Economic Review*, 2018, *108*, 775–827.

Aliprantis, Dionissi, "Redshirting, Compulsory Schooling Laws, and Educational Attainment," *Journal of Educational and Behavioral Statistics*, 2012, *37*.

Angrist, Joshua D. and Alan B. Keueger, "Does Compulsory School Attendance Affect Schooling and Earnings?," *The Quarterly Journal of Economics*, 11 1991, *106*, 979–1014.

Barua, Rashmi and Kevin Lang, "School entry, educational attainment, and quarter of birth: A cautionary tale of a local average treatment effect," *Journal of Human Capital*, 2016, *10*.

Bassok, Daphna and Sean F. Reardon, ""Academic Redshirting" in Kindergarten: Prevalence, Patterns, and Implications," *Educational Evaluation and Policy Analysis*, 2013, *35*.

Bedard, Kelly and Elizabeth Dhuey, "The persistence of early childhood maturity: International evidence of long-run age effects," *Quarterly Journal of Economics*, 2006, *121*.

Black, Sandra E., Erik Grönqvist, and Björn Öckert, "Born to lead? The effect of birth order on noncognitive abilities," *Review of Economics and Statistics*, 2018, *100*, 274–286.

Bleidorn, Wiebke, Ruben C. Arslan, Jaap J. A. Denissen, Peter J. Rentfrow,

Jochen E. Gebauer, Jeff Potter, and Samuel D. Gosling, "Age and gender differences in self-esteem—A cross-cultural window.," *Journal of Personality and Social Psychology*, 9 2016, *111*, 396–410.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43*, 972–1059.

_____, Bas ter Weel, and Bruce A. Weinberg, "Interpersonal Styles and Labor Market Outcomes," *Journal of Human Resources*, 2008, *43*, 815–858.

_____, _____, and _____, "Interpersonal Styles and Labor Market Outcomes," *Journal of Human Resources*, 2008, *43*, 815–858.

_____, _____, and _____, "People skills and the labor-market outcomes of underrepresented groups," *ILR Review*, 2014, *67*.

Bound, John, David A. Jaeger, and Regina M. Baker, "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 6 1995, *90*, 443.

Bowles, Samuel, Herbert Gintis, and Melissa Osborne, "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 2001, *39*, 1137–1176.

Braghieri, Luca, Ro'ee Levy, and Alexey Makarin, "Social Media and Mental Health," *American Economic Review*, 11 2022, *112*, 3660–3693.

Brooks-Gunn, Jeanne and Greg J. Duncan, "The Effects of Poverty on Children," *The Future of Children*, 1997, *7*, 55.

Buckles, Kasey S. and Daniel M. Hungerman, "Season of Birth and Later Outcomes: Old Questions, New Answers," *The Review of Economics and Statistics*, 7 2013, *95*, 711–724.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 11 2014, *82*, 2295–2326.

Cascio, Elizabeth U. and Diane Whitmore Schanzenbach, "First in the class? Age and the education production function," *Education Finance and Policy*, 2016, *11*.

_____ **and Ethan G. Lewis**, "Schooling and the Armed Forces Qualifying Test: Evidence from school-entry laws," *Journal of Human Resources*, 3 2006, *41*, 294–318.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 7 2020, *115*, 1449–1455.

Chao, Matthew and Euncheol Shin, "Cooperation and Coordination in Cultures with Age-Based Social Hierarchies," *SSRN Electronic Journal*, 2022.

Choi, Eleanor Jawon, Hyungsik Roger Moon, and Geert Ridder, "Estimation of an education production function under random assignment with selection," *American Economic Review*, 2014, *104*, 206–211.

Coley, Rebekah Levine, David Ribar, and Elizabeth Votruba-Drzal, "Do Children's Behavior Problems Limit Poor Women's Labor Market Success?," *Journal of Marriage and Family*, 2 2011, *73*, 33–45.

Cook, Philip J. and Songman Kang, "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*, 6 2020, *76*, 101968.

Cunha, Flavio and James Heckman, "The technology of skill formation," *American economic review*, 2007, *97* (2), 31–47.

_____ **and James J Heckman**, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 2008, *43*, 738–782.

_____, **James J. Heckman, and Susanne M. Schennach**, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78*, 883–931.

Datar, Ashlesha, "The impact of kindergarten entrance age policies on the childcare needs of families," *Journal of Policy Analysis and Management*, 12 2006, *25*, 129–153.

_____ **and Michael A. Gottfried**, "School entry age and children's social-behavioral skills: evidence from a national longitudinal study of U.S. kindergartners," *Educational Evaluation and Policy Analysis*, 2015, *37*.

de Araujo, Pedro and Stephen Lagos, "Self-esteem, education, and wages revisited," *Journal of Economic Psychology*, 2 2013, *34*, 120–132.

Dee, Thomas S. and Hans Henrik Sievertsen, "The gift of time? School starting age and mental health," *Health Economics (United Kingdom)*, 2018, *27*.

Deming, David J., "The growing importance of social skills in the labor market," *Quarterly Journal of Economics*, 11 2017, *132*, 1593–1640.

Dhuey, Elizabeth and Stephen Lipscomb, "What makes a leader? Relative age and high school leadership," *Economics of Education Review*, 2008, *27*.

_____, **David Figlio, Krzysztof Karbownik, and Jeffrey Roth**, "School Starting Age and Cognitive Development," *Journal of Policy Analysis and Management*, 2019, *38*.

Donati, Dante, Ruben Durante, Francesco Sobbrio, and Dijana Zejcirovic, "Lost in the Net? Broadband Internet and Youth Mental Health," 2022.

DuBois, David L., Robert D. Felner, Michelle D. Sherman, and Catherine A. Bull, "Socioenvironmental experiences, self-esteem, and emotional/behavioral problems in early adolescence," *American Journal of Community Psychology*, 6 1994, *22*, 371–397.

Dustmann, Christian, Hyejin Ku, and Do Won Kwak, "Why Are Single-Sex Schools Successful?," *Labour Economics*, 2018, *54*, 79–99.

Edin, Per Anders, Peter Fredriksson, Martin Nybom, and Björn Öckert, "The Rising Return to Noncognitive Skill†," *American Economic Journal: Applied Economics*,

2022, *14*.

Elder, Todd E., "The importance of relative standards in ADHD diagnoses: Evidence based on exact birth dates," *Journal of Health Economics*, 2010, *29*.

_____ **and Darren H. Lubotsky**, "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*, 2009, *44*.

Evans, William N., Melinda S. Morrill, and Stephen T. Parente, "Measuring inappropriate medical diagnosis and treatment in survey data: The case of ADHD among school-age children," *Journal of Health Economics*, 2010, *29*.

Fan, Jianqing and Irene Gijbels, *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, Vol. 66, CRC Press, 1996.

Fiorini, Mario and Katrien Stevens, "Scrutinizing the Monotonicity Assumption in IV and fuzzy RD designs*," *Oxford Bulletin of Economics and Statistics*, 2021, *83*.

Fletcher, Jason M., "Adolescent depression and educational attainment: results using sibling fixed effects," *Health Economics*, 7 2010, *19*, 855–871.

Fredriksson, Peter and Björn Öckert, "Life-cycle effects of age at school start," *Economic Journal*, 2014, *124*.

Gleason, Mary Margaret, Edward Goldson, Michael W. Yogman, Dina Lieser, Beth DelConte, Elaine Donoghue, Marian Earls, Danette Glassy, Terri McFadden, Alan Mendelsohn, Seth Scholer, Jennifer Takagishi, Douglas Vanderbilt, Patricia Gail Williams, Michael Yogman, Nerissa Bauer, Thresia B Gambon, Arthur Lavin, Keith M. Lemmon, Gerri Mattson, Jason Richard Rafferty, Lawrence Sagin Wissow, Carol Cohen Weitzman, Nerissa S. Bauer, David Omer Childers, Jack M. Levine, Ada Myriam Peralta-Carcelen, Peter Joseph Smith, Nathan J. Blum, Stephen H. Contompasis, Damon Russell Korb, Laura Joan McGuinn, and Robert G. Voigt, "Addressing Early Childhood

Emotional and Behavioral Problems," *Pediatrics*, 12 2016, *138*.

Golberstein, Ezra, Gilbert Gonzales, and Ellen Meara, "How do economic downturns affect the mental health of children? Evidence from the National Health Interview Survey," *Health Economics*, 8 2019, *28*, 955–970.

Gong, Jie, Yi Lu, and Hong Song, "The Effect of Teacher Gender on Students' Academic and Noncognitive Outcomes," *Journal of Labor Economics*, 2018, *36*, 743–778.

_____, _____, and _____, "Gender Peer Effects on Students' Academic and Noncognitive Outcomes Evidence and Mechanisms," *Journal of Human Resources*, 2021, *56* (3), 686–710.

Goodman, Robert, "The Strengths and Difficulties Questionnaire: A Research Note," *Journal of Child Psychology and Psychiatry*, 7 1997, *38*, 581–586.

Ha, Serin, "S Korea suicide: 'When I found my brother's body, my heart turned to winter'," December 2023. Accessed: 2025-05-01.

Hahn, Youjin, Liang Choon Wang, and Hee Seung Yang, "Does greater school autonomy make a difference? Evidence from a randomized natural experiment in South Korea," *Journal of Public Economics*, 2018, *161*, 15–30.

Hansen, Laura K. and Sara S. Jordan, "Internalizing Behaviors," *Encyclopedia of Personality and Individual Differences*, 2020, pp. 2343–2346.

Harter, Susan, "Causes and Consequences of Low Self-Esteem in Children and Adolescents," *Self-Esteem*, 1993, pp. 87–116.

Haushofer, Johannes and Jeremy Shapiro, "Erratum to "The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya"," *The Quarterly Journal of Economics*, 11 2017, *132*, 2057–2060.

Heckman, James J and Tim Kautz, "Hard evidence on soft skills," *Labour economics*, 2012, *19* (4), 451–464.

_____ and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the

GED Testing Program," *American Economic Review*, 2001, *91*, 145–149.

Heckman, James J., Jora Stixrud, and Sergio Urzua, "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior," *Journal of Labor Economics*, 2006, *24*, 411–482.

Hoeschler, Peter, Simone Balestra, and Uschi Backes-Gellner, "The development of non-cognitive skills in adolescence," *Economics Letters*, 2 2018, *163*, 40–45.

Huang, Francis L., "Investigating the Prevalence of Academic Redshirting Using Population-Level Data," *AERA Open*, 2015, *1.*

Jackson, C. Kirabo, "Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in Trinidad and Tobago," *Journal of Public Economics*, 2012, *96*, 173–187.

Jackson, C Kirabo, "What Do Test Scores Miss? The Importance of Teacher Effects on Non-Test Score Outcomes," *Journal of Political Economy*, 2018.

———, "Can Introducing Single-Sex Education into Low-Performing Schools Improve Academics, Arrests, and Teen Motherhood?," *Journal of Human Resources*, 2019.

Jackson, C. Kirabo, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastián Kiguel, "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment," *American Economic Review: Insights*, 12 2020, *2*, 491–508.

———, ———, ———, ———, **and** ———, "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment," *American Economic Review: Insights*, 2020, *2*, 491–508.

Johansen, Eva Rye, "Relative age for grade and adolescent risky health behavior," *Journal of Health Economics*, 2021, *76.*

Kauten, Rebecca and Christopher T. Barry, "Externalizing Behavior," *Encyclopedia of Personality and Individual Differences*, 2020, pp. 1509–1512.

Kim, Seonghoon, "The 6th-Year Study on the Development of Scale Scores for the Academic Achievement Test in the Seoul Education Longitudinal Study," Technical Report SERII 2015-42, Seoul Education Research & Information Institute, Seoul, Korea 2015.

Kim, Seul-Ki and Young-Chul Kim, "Non-Cognitive Consequences of Single-Sex Schooling," *Journal of Economic Theory and Econometrics*, 2022, *33* (2), 112–137.

Kim, Taehoon, "The Effect of Age at School Entry on Long-term Educational Attainment in Korea," *Korean Journal of Labor Economics*, 2011, *34* (1), 1–32.

_____, "Age culture, school-entry cutoff, and the choices of birth month and school-entry timing in South Korea," *Journal of Demographic Economics*, 2021, *87.*

Kraft, Matthew A., "Teacher effects on complex cognitive skills and social-emotional competencies," *Journal of Human Resources*, 2019, *54*, 1–36.

Kuhn, Peter and Catherine Weinberger, "Leadership skills and wages," *Journal of Labor Economics*, 2005, *23.*

Landersø, Rasmus, Helena Skyt Nielsen, and Marianne Simonsen, "School Starting Age and the Crime-age Profile," *Economic Journal*, 2017, *127.*

Landersø, Rasmus Kløve, Helena Skyt Nielsen, and Marianne Simonsen, "Effects of School Starting Age on the Family," *Journal of Human Resources*, 2020, *55.*

Larsen, Junilla K., Ad A. Vermulst, Rinie Geenen, Henriët van Middendorp, Tammy English, James J. Gross, Thao Ha, Catharine Evers, and Rutger C. M. E. Engels, "Emotion Regulation in Adolescence," *The Journal of Early Adolescence*, 2 2013, *33*, 184–200.

Lavy, Victor, "Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment," *Journal of public Economics*, 2008, *92* (10-11), 2083–2105.

_____, "Effects of Free Choice Among Public Schools," *Review of Economic Studies*, 2010, *77*, 1164–1191.

——— **and Edith Sand**, "The effect of social networks on students' academic and non-cognitive behavioural outcomes: Evidence from conditional random assignment of friends in school," *Economic Journal*, 2019, *129*, 439–480.

Lee, David S. and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, 2008, *142*.

Lee, Soohyung, Lesley J Turner, Seokjin Woo, and Kyunghee Kim, "All or Nothing? The Impact of School and Classroom Gender Composition on Effort and Academic Achievement," NBER Working Papers 20722, National Bureau of Economic Research 2014.

Lee, Valerie E. and Anthony S. Bryk, "Effects of Single-Sex Secondary Schools on Student Achievement and Attitudes," *Journal of Educational Psychology*, 1986, *78*.

Lee, Youngju and Nobuhiko Nakazawa, "Does Single-Sex Schooling Help or Hurt Labor Market Outcomes? Evidence from a Natural Experiment in South Korea," *SSRN Electronic Journal*, 2021.

Lim, Tae-Seop and Howard Giles, "Differences in American and Korean Evaluations of One-year Age Differences," *Journal of Multilingual and Multicultural Development*, 9 2007, *28*, 349–364.

Lindqvist, Erik and Roine Vestman, "The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment," *American Economic Journal: Applied Economics*, 2011, *3*, 101–128.

Liu, J., X. Chen, and G. Lewis, "Childhood internalizing behaviour: analysis and implications," *Journal of Psychiatric and Mental Health Nursing*, 12 2011, *18*, 884–894.

Lubotsky, Darren and Robert Kaestner, "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*, 2016, *53*.

Mael, Fred, Alex Alonso, Doug Gibson, Kelly Rogers, and Mark Smith, "Single-Sex Versus Coeducational Schooling : A Systematic Review," *Department of Education*,

2005.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 2008, *142.*

_____ **and Heather Royer**, "The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth," *American Economic Review*, 2 2011, *101*, 158–195.

McEachin, Andrew, Douglas Lee Lauen, Sarah Crittenden Fuller, and Rachel M. Perera, "Social returns to private choice? Effects of charter schools on behavioral outcomes, arrests, and civic participation," *Economics of Education Review*, 2020, *76.*

Mühlenweg, Andrea, Dorothea Blomeyer, Holger Stichnoth, and Manfred Laucht, "Effects of age at school entry (ASE) on the development of non-cognitive skills: Evidence from psychometric data," *Economics of Education Review*, 2012, *31.*

OECD, "Education at a Glance 2014: OECD Indicators," *OECD Publishing*, 2014.

Page, Lionel, Dipanwita Sarkar, and Juliana Silva-Goncalves, "Long-lasting effects of relative age at school," *Journal of Economic Behavior & Organization*, 12 2019, *168*, 166–195.

Pahlke, Erin, Janet Shibley Hyde, and Carlie M. Allison, "The effects of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis," *Psychological Bulletin*, 2014, *140.*

Park, Hyunjoon, Jere R. Behrman, and Jaesung Choi, "Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools," *Demography*, 2013, *50*, 447–469.

_____**,** _____**, and** _____, "Do single-sex schools enhance students' STEM (science, technology, engineering, and mathematics) outcomes?," *Economics of Education Review*, 2018, *62*, 35–47.

Patalay, Praveetha, Jay Belsky, Peter Fonagy, Panos Vostanis, Neil Humphrey,

Jessica Deighton, and Miranda Wolpert, "The Extent and Specificity of Relative Age Effects on Mental Health and Functioning in Early Adolescence," *Journal of Adolescent Health*, 11 2015, *57*, 475–481.

Pei, Zhuan, David S. Lee, David Card, and Andrea Weber, "Local Polynomial Order in Regression Discontinuity Designs," *Journal of Business and Economic Statistics*, 2022, *40*, 1259–1267.

Persson, Petra and Maya Rossin-Slater, "Family Ruptures, Stress, and the Mental Health of the Next Generation," *American Economic Review*, 4 2018, *108*, 1214–1252.

Peña, Pablo A., "Relative age and investment in human capital," *Economics of Education Review*, 2020, *78*.

_____ **and Angela L. Duckworth**, "The effects of relative and absolute age in the measurement of grit from 9th to 12th grade," *Economics of Education Review*, 2018, *66*.

Pintrich, Paul R., David A. F. Smith, Teresa Garcia, and Wilbert J. Mckeachie, "Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq)," *Educational and Psychological Measurement*, 1993, *53*.

Richard, Patrick, Darrell J. Gaskin, Pierre K. Alexandre, Laura S. Burke, and Mustafa Younis, "Children's Emotional and Behavioral Problems and Their Mothers' Labor Supply," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 1 2014, *51*, 004695801455794.

Ridley, Matthew, Gautam Rao, Frank Schilbach, and Vikram Patel, "Poverty, depression, and anxiety: Causal evidence and mechanisms," *Science*, 12 2020, *370*.

Riordan, Cornelius, "Public and Catholic Schooling: The Effects of Gender Context Policy," *American Journal of Education*, 1985, *93*.

Shigeoka, Hitoshi, "SCHOOL ENTRY CUTOFF DATE AND THE TIMING OF BIRTHS," 2015.

Shin, Estelle, "The Persistent Effect of Being Old for Grade on Social-Emotional Skills,"

2023.

Snyder, Thomas D and Sally A Dillow, "Digest of Education Statistics, 2012," Technical Report NCES 2014-015, National Center for Education Statistics, Washington, DC 2013.

Sohn, Hosung, "Mean and distributional impact of single-sex high schools on students' cognitive achievement, major choice, and test-taking behavior: Evidence from a random assignment policy in Seoul, Korea," *Economics of Education Review*, 2016, *52*, 155–175.

Sorrenti, Giuseppe, Ulf Zölitz, Denis Ribeaud, and Manuel Eisner, "The causal impact of socio-emotional skills training on educational success," *Review of Economic Studies*, 2025, *92* (1), 506–552.

Sowislo, Julia Friederike and Ulrich Orth, "Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies.," *Psychological Bulletin*, 1 2013, *139*, 213–240.

Steger, Michael F, Patricia Frazier, Shigehiro Oishi, and Matthew Kaler, "The meaning in life questionnaire: assessing the presence of and search for meaning in life.," *Journal of counseling psychology*, 2006, *53* (1), 80.

Weidmann, Ben and David J Deming, "Team players: How social skills improve team performance," *Econometrica*, 2021, *89* (6), 2637–2657.

Weinberger, Catherine J., "The Increasing Complementarity between Cognitive and Social Skills," *Review of Economics and Statistics*, 2014, *96*.