The Asymmetric Effects of High Achiever Peers: Experimental Evidence from Ecuador*

Nicolás Fuertes-Segura† Yyannú Cruz-Aguayo Carolina Echeverri

January 25, 2025

Abstract

We study the impact of exposure to high achievers, identified through teacher rankings of students' performance, on cognitive and non-cognitive skills using a unique randomized experiment from Ecuador. Students in each school are randomly assigned to classrooms in every grade from 1st to 6th. The proportion of high-achieving peers varies across grades and classrooms due to the random assignment. We find that exposure to high achievers reduces test scores in math and executive function. Additionally, male students influence only their male peers, while female students affect only female peers. Competition within the classroom intensifies peer effects, with the impact of high achievers concentrated among students in the top quintile of the previous year's test scores and those attending smaller schools. Finally, as with other school inputs, the effects of 1st-grade peers are stronger but fade over time. We also find reductions in self-reported happiness, but no impacts on non-cognitive skills.

Keywords: High Achievers, Peer effects, Competition, Elementary School

JEL: I21, J16

*Fuertes-Segura: UC Santa Barbara; Cruz-Aguayo: Inter-American Development Bank; Echeverri: University of Minnesota

†Corresponding author. Mailing Address: Department of Economics, UC Santa Barbara. North Hall 3017, Santa Barbara, CA 93117. E-mail: nfuertessegura@ucsb.edu

1 Introduction

Peer effects significantly shape individual outcomes across various settings, including education, the workplace, and social environments. These effects influence behaviors, performance, and decision-making processes. Previous studies have demonstrated that peer composition impacts students' outcomes through multiple mechanisms, primarily in developed countries. Factors such as classroom size (Hoxby, 2000b), peers' achievement (Hanushek et al., 2003) and gender (Hoxby, 2000a), sorting into friends pairings (Card and Giuliano, 2013), and rank effects (Murphy and Weinhardt, 2020) have been studied. However, there is limited knowledge about the effects of peer composition on educational outcomes in developing countries where the education systems face different incentives and resources are more limited.

This paper investigates the impact of the proportion of high achievers, identified through teacher rankings of students' performance, on both cognitive (math and executive function) and non-cognitive skills (depression, self-esteem, growth mindset, and grit) throughout elementary school. We use data from a unique randomized experiment involving the entering Kindergarten cohort (over 20,000 students) across 202 elementary schools in Ecuador in 2012. For the same cohort, we have rich data on baseline information, test scores, teachers' information, and classroom quality information for seven consecutive years. At the beginning of each grade, from Kindergarten to 6th grade, students were randomly assigned to classrooms within their school, with every school having at least two classrooms per grade. Compliance with the random assignment was nearly perfect (98.9%, on average over the seven years of the experiment). Consequently, children who did not switch schools were exposed to seven exogenous, orthogonal sets of peers, and variations in peer groups are attributable to this random assignment, leading to different proportions of high achievers in each classroom. This feature enables us to provide a causal interpretation of the peer effects of high achievers.[1]

The impact of high achievers on their peers' outcomes is context-dependent, with evi-

[1]Given the design, there is variation in classroom and teacher quality. However, these factors are orthogonal to peer quality as teachers were also assigned randomly to classrooms within schools and grades.

dence yielding mixed results. On the positive side, high achievers can enhance their peers' performance by fostering positive peer pressure and raising academic standards. This exposure has been shown to improve students' achievement (Balestra et al., 2023), aspirations, parental investments (de Gendre and Salamanca, 2020), and ultimately their lifetime earnings (Bertoni et al., 2020). Conversely, the presence of high-achieving peers may foster competition, which could lead to negative effects. Theoretical research suggests that competition could discourage collaborative behaviors, such as helping efforts, while increasing individual effort (Drago and Garvey, 1998). These competitive dynamics may not affect all students equally. For example, evidence shows differences between genders in effectiveness in competitive environments (Gneezy et al., 2003). As a result, high-achieving peers might negatively impact certain groups, such as high-achieving girls (Cools et al., 2022) and privileged students from better-educated families (Bertoni et al., 2020). This illustrates the nuanced nature of peer influence and shows how peer effects have strong implications for human capital accumulation and gender gaps.

In the main specification, we include school-by-grade fixed effects, which allow us to compare children who attended the same school and grade but were randomly assigned to different classrooms, thus exposing them to varying proportions of high achievers. At the end of each grade, we asked teachers to identify five students with the highest performance. In this specification, a child is considered a high achiever only if at least 50% of the teachers who observed them in previous grades identified them as such. We find that a one standard deviation increase in the leave-one-out proportion of high achievers reduces math test scores by 0.011 SD and executive function skills—a set of basic self-regulatory skills—by 0.014 SD. These results are surprising given the previous literature on high achievers that found mostly positive effects but are consistent with those found by Hoxby and Weingarth (2005). Regarding the magnitude, the effects are slightly smaller than the 0.04SD decrease in GPA found by Chen and Hu (ming). However, their estimate is for high-ability students instead of all the children in the classroom, which could explain the differences. Alongside these reductions in cognitive skills, we find no effect on non-cognitive skills in the 6th grade and a decrease in self-reported happiness in the 1st grade.

The composition of the peer group and the individual characteristics significantly impact educational outcomes. Gender dynamics, in particular, can influence how students respond to peer composition, with men and women exhibiting distinct reactions (Feld and Zölitz, 2017; Cools et al., 2022; Busso and Frisancho, 2021). Therefore, when separating the results by gender, we observe that male high achievers only impact male students, whereas female high achievers impact only female students. Notably, these results are unexpected as previous results find that male students negatively affect female ones (Busso and Frisancho, 2021; Cools et al., 2022). However, elementary school students tend to befriend more students of the same gender (McPherson et al., 2001; Garrote et al., 2023), which could explain why these results differ from those found previously. Indeed, in our sample, 89.6% of the males report having a male best friend and 92.7% of the females report having a female best friend. Importantly, these results are consistent with previous evidence of how males are negatively affected by their high-achieving male peers (Lavy et al., 2012). This highlights how peer and gender composition affect cognitive skills during elementary school.

A second set of results analyzes how student competition influences peer effects. Specifically, competition is expected to impact students in the top quintile of the distribution more strongly, as they are more likely to receive recognition for their work. As anticipated, we find that high-achieving classmates have detrimental effects on the academic performance of high-performing students. We find that a one standard deviation increase in the proportion of high achievers reduces math test scores by 0.030 SD for high-perfoming children. The magnitude of our effects is similar to the 0.040 SD reduction in GPA found for a comparable sample of students by (Chen and Hu, ming). Additionally, declines in happiness appear to be concentrated among these top-performing students. Remarkably, these results suggest that students who were top performers in the previous year may experience a stronger negative motivational impact when more high achievers join their classes, potentially impacting their test scores.

Additionally, we uncover how these negative effects increase with the intensity of competition within a given school. First, we examine how the effects vary by the number of high achievers in a classroom, as competition is expected to be more pronounced when there are more high achievers to compete with. Consistent with this, our results show that the negative effects are largest in classrooms with a higher number of high achievers (on average, eight). Second, incentives to compete decrease with increasing school size. For instance, in schools with an honor roll, the expected benefits of surpassing classmates diminish as school size grows. Conversely, in smaller schools, the recognition of being a high achiever is more salient, as it is clearer who holds that status. Indeed, we find that school size significantly mitigates negative peer effects, with detrimental impacts mainly in smaller schools. Overall, our analysis of competition across varying intensities suggests that negative effects are more pronounced in competitive environments. Furthermore, we explore how these two dimensions interact, finding that the negative effects of having more high achievers are stronger in smaller schools.

We analyze how teacher quality, assessed by observed teacher behaviors, mediates the peer effects. We find that higher-quality teachers, particularly those skilled in classroom organization, can mitigate the detrimental effects on learning. These results suggest that teacher quality plays a crucial role in lessening the negative impact of high achievers, especially through effective classroom organization and group management. Moreover, the findings underscore the need for further investigation into how peer effects develop in classrooms with low-quality teachers and the specific roles these teachers play in shaping outcomes.

Finally, a novel aspect of our paper is exploring how peer composition effects evolve over time. We find that peer effects are most pronounced in the early grades, particularly in 1st grade. These results align with previous research on other school and classroom inputs, which suggests that younger children may be more sensitive to environmental influences than older children, even within elementary school (Aizer, 2008; Carneiro et al., ming). Furthermore, we observe that the initial impact in 1st grade declines significantly as children age. After five years, the 1st-grade effect is reduced to nearly a third of its original size. These findings are consistent with existing literature showing that the effects of other achievement determinants, such as teacher quality or value-added, also tend to fade over time (Jacob et al., 2010; Chetty et al., 2014b).

This paper contributes to the literature in four ways. First, it contributes to the literature on the experimental evidence of peers' composition effects in learning environments. While previous papers have studied the random assignment of peers in middle schools (Busso and Frisancho, 2021), college settings (Sacerdote, 2001; Carrell et al., 2009; Carrelll et al., 2013; Sacerdote, 2014; Feld and Zölitz, 2022), this study directly focuses on how peer composition impacts educational outcomes during elementary school, a period in which children's environment could have potential long-term effects (Heckman et al., 2013). We use a unique randomized experiment with multiple rounds of random assignment, and essentially perfect compliance.

Second, this paper directly contributes to the literature on the impacts of high-achievers. While previous studies have examined the effects of high achievers on test or admission scores (Busso and Frisancho, 2021; Cools et al., 2022) and college entrance outcomes (Cools et al., 2022; Mouganie and Wang, 2020), evidence regarding their impact on executive function and non-cognitive skills is limited. This is significant because previous evidence has shown that executive function and non-cognitive skills develop during childhood and can influence adult outcomes (Moffitt et al., 2011; Heckman et al., 2006). More importantly, in line with prior findings, we demonstrate that these impacts vary with competition intensity, even in lower-stakes environments such as elementary schools (Chen and Hu, ming). Finally, our study benefits from rich data on teacher perceptions and classroom activities compared to previous research, allowing us to examine the formation of multiple skills in elementary school and how these skills are influenced by peer composition, competition, and their interaction with other classroom inputs. This highlights potential avenues for policy design.

Third, we contribute to the literature exploring the timing and the persistence of the effects of various determinants of achievement (Jacob et al., 2010; Chetty et al., 2014b; Carneiro et al., ming). We expand on this literature by examining whether peer composition affects cognitive and non-cognitive skills differently across grades. Furthermore, given the longitudinal nature of the data, we can estimate how the effects evolve over time and test whether our findings are consistent with previous research on other school inputs (Jacob et al., 2010; Chetty et al., 2014b).

Finally, even though peer effects are context-specific, little is known about how peer composition affects educational outcomes outside of developed countries.[2] We provide evidence on peer effects in a new setting: elementary schools in Ecuador (Mouganie and Wang, 2020; Busso and Frisancho, 2021; Chen and Hu, ming). Elementary schools in developing countries face different incentives, tighter resource constraints, and larger classroom sizes. This study highlights that policies aimed at improving students' outcomes must consider the differences in peer effects across countries.

The remainder of the paper is organized as follows. Information about Ecuador, the experimental setting, and data can be found in section 2. Section 3 outlines the empirical strategy. Section 4 provides the main results, the dynamics of the effects, and heterogeneity analysis. Conclusions and policy discussion can be found in section 5.

2 Setting and Data

2.1 Education system in Ecuador

Ecuador is a middle-income country and one of the smallest in South America, with a population of 15.7 million and a GDP per capita of approximately $11.300 (in PPP U.S. dollars) in 2013. Schooling is compulsory from 5 to 14 years of age. The education system is divided into elementary school (Kindergarten through 6th grade), middle school (7th through

[2]See Sacerdote (2011, 2014) for a review of previous studies

9th grades), and high school (10th through 12th grades). The school year follows a dual-calendar system in which the coastal region runs from May to February (similar to most countries in the Southern Hemisphere and many in Latin America), and from September to June in the highlands and Oriente (eastern) regions (similar to most countries in the Northern Hemisphere)[3]

Approximately 4.4 million children were enrolled in the education system during the 2012–2013 school year, with 78 percent attending public schools.[4] Ecuador has made significant progress in expanding the coverage of its education system. In 2013, the primary school net enrollment rate was 95%, with a 97% completion rates.[5]. However, math achievement among young children remains low (Berlinski and Schady, 2015; Näslund-Hadley and Bando, 2015). For instance, in 2013, only 48.4% of the country's children in primary school achieved the minimum proficiency level in mathematics, with a slight gender gap—49.8% for males compared to 46.7% for females.[6] Therefore, like many Latin American nations, Ecuador's critical educational challenge is quality, not access.

The system employs more than 208,000 teachers in the public sector, with salaries primarily determined by seniority. Around 53% of teachers in Ecuador hold tenure, while the remaining 47% work on a contract basis. Between 2012 and 2019, the proportion of qualified teachers in primary increased from 80.1% to 89.3%, contributing to improvements in the quality of the education system and the instruction provided to students.[7]. However, key challenges persist; for example, the pupil-qualified teacher ratio remained stable at around 27.5 during this period.[8]

[3]According to the 2010 census, 53% of the population lived in the coast, 42% lived in the highlands, and 5% lived in the Oriente region.

[4]Data obtained from https://educacion.gob.ec/datos-abiertos/ on November 4th, 2024

[5]The net enrolment rate is the fraction of children of school age who are enrolled in school, while the completion rate corresponds to the percentage of children that have finished the last grade of primary.

[6]It corresponds to the percentage of children in primary that exceed the minimum proficiency level (MPL) which is the pre-defined proficiency level of basic knowledge in mathematics measured through learning assessments. All data was obtained from the World Development Indicators (https://datatopics.worldbank.org/world-development-indicators/) on November 4th, 2024

[7]It corresponds to the fraction of teachers with the minimum required academic qualifications required for teaching in primary school.

[8]All data was obtained from the World Development Indicators (https://datatopics.worldbank.org/world-

2.2 Experimental setting

We use a unique experiment conducted in 202 schools in Ecuador, representative of the country's public education system in the coastal region, to study how peer composition impacts cognitive and non-cognitive skills.[9] Each school had at least two classrooms per grade (most had exactly two). In the 2012 school year, an incoming cohort of children was randomly assigned to Kindergarten classrooms within each school. These children were randomly reassigned to higher-grade classrooms in subsequent years until 6th grade in 2018. Compliance with the assignment rules was very high, averaging 98.9%. Thus, children who remained in the same school for the entirety of the elementary school cycle were exposed to seven exogenous, orthogonal shocks to classroom quality.

The random assignment allows us to address concerns about the purposeful matching of students with teachers and peers, a common issue in non-experimental settings (Chetty et al., 2014a; Rothstein, 2010).[10] To test the success of the random assignment, we apply the method proposed by Jochmans (2023), which checks for correlation between student i's end-of-grade scores in year $t-1$ and the average scores in $t-1$ of the classroom peers assigned to her in year t. As shown in Appendix Table B.1, we cannot reject the null hypothesis that there is no correlation between child i's achievement and that of her classroom peers, confirming that random assignment was successful in this context. Further details on the classroom assignment rules, randomization tests, and compliance with randomization are provided in Appendix B.

development-indicators/

[9]These schools are a random sample of all public schools with at least two Kindergarten classrooms in the country's coastal region 2012. See Araujo et al. (2016) for more details on the selection of the schools. Morevover, they show that the characteristics of students and teachers in the sample are very similar to those of students and teachers in a nationally representative sample of schools in Ecuador.

[10]We use the term "random" for simplicity, but strictly speaking, random assignment only occurred from 3rd to 6th grade. In the other grades, the assignment rules were as-good-as-random. Specifically, in Kindergarten, all children in each school were ordered by their last and first names, then assigned to teachers in alternating order. In 1st grade, they were ordered by date of birth, from oldest to youngest, and assigned to teachers in alternating order. In 2nd grade, they were divided by gender, ordered by first and last names, and then assigned in alternating order. From 3rd to 6th grades, students were separated by gender and randomly assigned to classrooms.

2.3 Data

2.3.1 Child data

At the beginning of Kindergarten, we collected baseline data on maternal education, household wealth, preschool attendance, and children's vocabulary skills.[11] We administered age-appropriate math tests at the end of each grade from Kindergarten to 6th grade. These tests included material that teachers were expected to have covered explicitly in class (e.g., addition or subtraction), material that may have been covered during the academic year but likely in a somewhat different format (e.g., simple word problems), and material that was not covered in class but has been shown to predict current and future math achievement (e.g., the Siegler number line task).[12] We aggregated correct math responses for each component separately using Item Response Theory (IRT) and then computed the total math score where each component receives the same weight.

Additionally, we collected test data on executive function for each grade from Kindergarten to 4th grade. Executive function encompasses a set of basic self-regulatory skills involving various regions of the brain, particularly the prefrontal cortex. Executive function is generally divided into three domains: working memory, inhibitory control, and cognitive flexibility.[13] These skills are crucial for young children to adapt and learn effectively in school, as they are needed to pay attention, take turns, ask questions, remember steps, and

[11] To measure baseline receptive vocabulary, we use the Test de Vocabulario en Imágenes Peabody (TVIP), the Spanish-speaking version of the widely-used Peabody Picture Vocabulary Test (PPVT) (Dunn et al., 2015). The TVIP was normed on samples of Mexican and Puerto Rican children and has been widely employed to assess verbal ability and development among Latin American children.

[12] The number line task works as follows: children are shown a line with endpoints marked. For example, in 1st grade, the left end of the line is marked with a 0, and the right end is marked with a 20. They are then asked to place various numbers on the line (e.g., 2 or 18). The accuracy with which children place the numbers has been shown to predict general math achievement (see Siegler and Booth (2004)).

[13] Working memory measures the ability to retain and manipulate information. For example, 2nd-grade children were asked to remember increasingly long strings of numbers and repeat them in order and then backward. Cognitive flexibility measures measures the ability to shift attention between tasks and adapt to different rules. For example, 1st-grade children were shown picture cards featuring trucks or stars, in red or blue, and were first asked to sort the cards by *shape* (trucks versus stars) and then by *color* (red versus blue). Inhibitory control refers to the capacity to suppress impulsive responses. For example, Kindergarten children were quickly shown a series of flashcards displaying either a sun or a moon and were asked to say "day" when they saw the moon and "night" when they saw the sun.

solve math problems, among other classroom tasks. More importantly, these skills have been shown to predict outcomes into adulthood related to labor market, health and crime (Moffitt et al., 2011). We compute scores for each of the three domains in executive function, as well as an overall score.[14]

In 1st grade, we collected measures of happiness and effort. For happiness, we asked children if they were always, sometimes, or never happy at school and in their classroom. Similarly, we asked about effort, inquiring if they always, sometimes, or never made an effort to understand what the teachers were teaching and to learn as much as possible in school. Most children in the sample answered the same option for both questions related to happiness and effort, respectively. We aggregated the responses for each variable to construct two self-reported dummy variables: (i) children who were always happy and (ii) children who always put in effort.

In 6th grade, we collected data on child depression, self-esteem, growth mindset, and grit. To measure child depression, we used the Patient-Reported Outcomes Measurement Information System (PROMIS) Depression Scale for children aged 11-17 years, developed by the American Psychiatric Association. We selected five questions from the National Longitudinal Study of Adolescent to Adult Health (Add Health) to measure self-esteem. For growth mindset, we selected 10 of the 20 questions on the Dweck "Mindset Quiz"; growth mindset refers to the belief that intelligence is malleable, rather than fixed, and can be increased with effort (Blackwell et al., 2007). Lastly, to measure grit, we adapted four questions from the 8-item Grit Scale for children (Duckworth and Quinn, 2009); grit refers to the capacity of individuals to persevere at a given task. For each of these outcomes, we aggregate responses by factor analysis to aggregate responses for each outcome. Further details on all the child assessments and the IRT are provided in Appendix C.

[14]Unlike math test scores, we cannot aggregate questions using factor analysis since some tests are timed. For instance, in one task, children have 2 minutes to find as many sequences of "dog", "house", and "ball" in that exact order on a sheet with rows of dogs, houses, and balls in various possible sequences. The score on this test is the number of correct sequences the child finds. Therefore, we calculate separate scores for working memory, inhibitory control, cognitive flexibility, and a total executive function score, with each domain given equal weight.

Most of the tests were administered to children individually by specially trained enumerators. The only exception is some tests in 4th through 6th grades, which were applied in a group setting. All tests, except for the non-cognitive tests applied in 6th grade, were conducted at school. Before the official testing, a pilot study was conducted to select appropriate questions based on difficulty and discrimination parameters obtained from an Item Response Theory (IRT) procedure. We chose questions that were easily comprehensible to the children based on their context and exhibited reasonable variability in the pilot to capture the complete distribution of students' abilities.

2.3.2 High Achievers Information

At the end of each grade, we asked teachers to identify five students with the highest performance.[15] The teachers were not provided access to any test results from the experiment and had to rely solely on the information they knew about the students. Based on this information, we calculated the proportion of peers who are high achievers in each classroom. However, since teachers' rankings are subjective, we considered a child a high achiever only if at least 50% of the teachers who observed them in previous grades identified them as such.[16] This implies that what constitutes a high achiever may vary over time. As children grow, more teachers can assess their performance, and they could stop being considered high achievers or start being considered high achievers at any grade.

There are several ways to identify high-achieving students. For instance, one could use measures of innate ability (e.g., IQ scores), baseline characteristics (e.g., PPVT scores), parents' education (Cools et al., 2022), or results from placement exams (Busso and Frisancho, 2021). However, we prefer to rely on teachers' rankings to define high achievers for three main reasons. First, this approach enables us to identify high achievers each year, which other measures do not consistently allow. While baseline scores may provide valuable in-

[15] Importantly, we did not ask teachers who they thought were top and bottom performers in math and, separately, in language.

[16] For example, a child in 4th grade identified as a top performer by three former teachers would be considered a high achiever for that year.

sights, we do not have information for all the children in the sample. For example, PPVT scores, though informative, are only available for roughly 14,000 of the nearly 28,000 students in our sample. Second, due to their design, test scores in the next grade combine previously asked and new questions. This design feature implies that test scores are correlated over time, which could lead to mechanically classifying students as high achievers across multiple years. Finally, teachers' rankings reflect classroom performance but do not capture performance on our test scores, allowing them to serve as a distinct measure of "ability". Teachers are also likely to value other characteristics beyond academic performance (e.g., persistence, grit) that would not be captured by test scores or by measures based on parents' education.

2.3.3 Teacher and Classroom data

We use the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2015) to measure teacher behaviors. The CLASS measures teacher behaviors in three domains: Emotional Support, Classroom Organization, and Instructional Support.[17] Within each of these domains, there are 3 to 4 CLASS dimensions that are rated on a scale of 1 to 7. For each behavior, the CLASS protocol provides coders with concrete guidance on whether the score should be categorized as "low" (scores of 1–2), "medium" (3–5), or "high" (6–7). The behaviors that coders are looking for in each dimension are quite specific.

The CLASS has been widely used for research and policy purposes in the U.S., especially in preschool settings. For instance, Head Start grantees in the U.S. are required to achieve a minimum score on the CLASS to be re-certified for funding. The CLASS has also been employed as a measure of teacher quality in Latin America, including earlier work in Ecuador Araujo et al. (2016), Chile (Bassi et al., 2020; Yoshikawa et al., 2015) and Peru (Araujo et al., 2019). All Kindergarten through 4th-grade teachers were filmed for a full day (from approximately 8 a.m. to 1 p.m.) without prior notice. Teachers were only informed of

[17] Emotional Support measures how teachers promote a positive climate in their classroom and respond to students' emotional needs. Classroom Organization refers to the routines and procedures established by the teacher to manage students' behavior, time, and attention. Finally, Instructional Support evaluates how effectively a teacher promotes students' thinking by implementing the curriculum.

the filming day on the day itself, ensuring that the video captured their typical, planned interactions with the students. Each classroom has only one teacher, without a classroom aide, who is responsible for all academic subjects (all subjects other than physical education and art and music when available). We strictly adhered to CLASS protocols when coding the footage. Each video recording was divided into 20-minute segments, with two coders evaluating each segment.[18] Further details on the CLASS dimension, the scoring protocol, and its application in Ecuador are provided in Appendix D.

One disadvantage of the CLASS is that it is not a cardinal scale, which implies that it cannot be use as an input in the production function, unlike value-added, experience, and other teacher characteristics. However, previous studies have found that teacher behaviors, measured by the CLASS, are associated with higher test scores (Araujo et al., 2016). Therefore, following Araujo et al. (2016), we use the CLASS as a measure of Responsive Teaching. We divide the teachers into two groups (low and high quality) based on their CLASS score, separately for each component and the total score.

2.4 Descriptive Statistics and Preliminary Checks

Table 1, Panel A provides summary statistics for the children in the sample. The table shows that, on average, children were five years old on the first day of Kindergarten, with 60% having attended preschool, and half being girls. Additionally, the average age of the children's parents is 30 years, and they have completed only eight years of education. The average score on the vocabulary skills (PPVT) test is 83 points, with a considerable range from 55 to 145. Furthermore, 90% of the children self-report being happy, and 85% of the children self-report putting in effort. More importantly, 13% of the children are classified as high achievers by their previous teachers. Table 1, Panel B shows information about the household and parents' characteristics of the children in the sample. On average, the parent's completed approximately 8.5 years of schooling.

[18] In cases where the two coders provided significantly different scores, a third coder was assigned to evaluate the teacher.

Table 1, Panels C and D show the Classroom and Teacher characteristics, respectively. The panels indicate that, on average, 15% of the peers of a student in a given classroom are high achievers and that the average classroom contains 37 students. However, these proportions vary significantly across classrooms (with a range from 0 to 57%). The average teacher in the sample has 18 years of experience. Additionally, 80% are tenured, and 84% are women. Importantly, teacher quality, as measured by the CLASS, varies across classrooms in every assessment domain.

Before estimating the impacts of high achievers on cognitive and non-cognitive skills, we provide descriptive evidence regarding the variation in the proportion of high achievers, their characteristics, and potential estimates. First, the experiment generates considerable variation in exposure to high achievers. Appendix Figure A.1 illustrates the distribution of the proportion of high achievers in classrooms across all grades. Although only 13% of the children in the sample are classified as high achievers, most of them are distributed across different classrooms. Indeed, more than 60% of the classrooms had three or more achievers. As a result, only 26 classrooms did not contain a high achiever.[19]

Second, one potential concern with our preferred measure of high achievers is whether teachers can effectively identify them. Table 2 provides summary statistics of the characteristics of the children in our sample, comparing those classified as high achievers based on the teacher's rankings with those not. The table shows that children classified as high achievers obtained high math and executive function scores, as well as the baseline vocabulary test (PPVT). High achievers also report higher levels of self-reported happiness and effort. Overall, there are notable differences in socioeconomic status between high achievers and others; for instance, high achievers have more educated parents and are more likely to have attended preschool. There are no significant differences in the remaining socioeconomic variables. This suggests that teachers appear to identify the high achievers correctly, at least based on their observable characteristics, while also considering other relevant characteristics.

[19] This corresponds to 17 classrooms in 1st grade, 0 in 2nd grade, 1 in 3rd grade, 1 in 4th grade, 5 in 5th grade, and 2 in 6th grade.

Finally, Figure 1 illustrates the relationship between the (leave-one-out) proportion of high achievers and math test scores, conditional on the school-by-grade fixed effects. Specifically, it compares students who attend the same school but were exposed to different proportions of high achievers due to the random assignment without considering other factors like ability or gender. The figure suggest that children exposed to more high achievers obtained lower math test scores.

3 Empirical Strategy

3.1 Main Model

The goal is to estimate the impact of a child's peers' composition on her subsequent learning in elementary school. The dataset allows us to construct measures of the (leave-one-out) proportion of high, lagged achievement, and achievement at the current grade for children between 1st grade and 6th grade. With these data, we can investigate the impact of peer composition in the short- and medium-run. We use unique data on rankings provided by the teachers. As we explained previously, a child is considered a high achiever only if at least 50% of the teachers who observed them in previous grades identified them as such. Therefore, to measure the impact of high achievers on cognitive and non-cognitive skills, we estimate the following model pooling observations from all grades:

$$y_{icst} = \beta_1 FracHighAchievers_{cst} + \mathbf{X}_{icst}\alpha + \delta_{st} + \varepsilon_{icst} \tag{1}$$

where y_{icst} is one of students' i performance (measured by the corresponding test score) in classroom c in school s at the end of grade t . $FracHighAchievers_{cts}$ is the leave-one-out proportion of high achievers peers when the student is randomly assigned to classroom c in school s at grade t. This variable equals the proportion of students who are high achievers from the school-grade-classroom distribution of students after eliminating student i from the distribution. δ_{st} is a school-by-grade fixed effects. $\mathbf{X}_{icst}$ is a vector of individual-level controls such as children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. We cluster standard errors at the school-by-grade level, which allows students' outcomes in different classrooms to correlate within schools on a given grade. The parameter of interest, β_1, measures the impact of the (leave-one-out) proportion of high-achievers on the test scores.

Previous literature has found that how girls and boys react to high achievers depends on whether these high achievers are girls or boys (Busso and Frisancho, 2021; Cools et al., 2022). We explore this question by analyzing how girls' and boys' test scores differ when faced with high-achieving peers, separated by gender. Therefore, the model we estimate is the following:

$$y_{icst} = \beta_1 MaleFracAchievers_{cst} + \beta_2 FemaleFracAchievers_{cst} + \mathbf{X}_{icst}\alpha + \delta_{st} + \varepsilon_{icst} \quad (2)$$

where $MaleFracAchievers_{sct}$ ($FemaleFracAchievers_{sct}$) are the leave-one-out proportion of high achievers male (female) peers in classroom c at grade t in school s. Specifically, they are the proportion of students from the specific gender-school-grade-classroom distribution of students after eliminating student i from it. The parameters of interest, β_1 and β_2, measure the impact of the (leave-one-out) proportion of male and female high-achievers on the test scores, respectively

Identification

The empirical strategy exploits variation in exposure to high achievers across different classrooms in elementary school students within a given school and grade, which is a common approach in the literature. The key identification assumption is that conditional on school-by-grade fixed effects, classroom with and without high achievers do not differ systematically in any other dimension. The school-by-grade fixed effects control for differences across cohorts within a given school and aspects of the school-grade that are constant across students. Moreover, intuitively, this key assumption holds because of the random assignment of students to classrooms every grade.

We provide two tests for random assignment. First, the ability to exploit peer composition relies on sufficient residual variation in the main variables after considering the set of fixed effects. Given the random assignment, we would expect that the corresponding distribution of the proportion of high achievers in a given classroom to look normally distributed, conditional on school-by-grade fixed effects. Figure 2 shows that, as expected, the corresponding distribution of the proportion of high achievers indeed looks normally distributed. This figure suggests that peer composition is likely arbitrary.

Second, we test if the variation in the proportion of high achievers within a school-grade is consistent with a random assignment process by comparing the actual distribution to a simulated distribution of the proportion of high achivers. To this end, we do Monte Carlo simulations in which we randomly assign students to classroom within the same school and grade. Similarly to Bietenbeck (2020), we take the number and size of classrooms and the number of high achievers in the classroom from the actual data. In the simulated data, we regress the proportion of high achievers on school-by-grade fixed effects and collect the residuals. In Figure A.2, we plot the distribution of the residuals from 1,000 replications and the residuals from the actual data. The two distributions look similar which is consistent with the random assignment that took place and further supports the assumption that high achievers were randomly assign to classrooms within school-grades.

Nevertheless, to further support the causal interpretation of the estimates, we provide several sensitivity tests. First, sensitivity tests are conducted for different ways of controlling for differences across schools and grades and to address any concerns regarding omitted factors influencing the results. Specifically, we estimate three alternate specifications: i) we exclude the grade fixed effect; ii) we separate the school and grade fixed effects; (iii) we include teacher controls to account for how teachers could change the classroom environment and affect the students outcomes. As we show later, these results are similar to those obtained using the main specification.

Second, we test the sensitivity of the estimates to different ways of defining who is a high achiever in the classroom. Regarding our preferred measure, we estimate three new proportions by defining a child as high achiever only if at least 25%, 75%, or 100% of the teachers who observed them in previous grades identified them as such instead of the original 50%. We also create a new measure of high achiever that uses only the information available for the previous grade teacher instead of using the information from all teachers. Given the availability of test scores, we also construct two measures based on the results of the children: (i) whether a student is over the 95th percentile of the corresponding test score the previous year, and (ii) whether a student is over the 95th percentile of the baseline PPVT test. The results do not show significant changes.[20]

Third, we test the sensitivity of the estimates to the exclusion of one school at a time to ensure that the results are not driven by a single school in a given grade. Appendix Figure A.3 show the distribution of the leave-one-out coefficient estimates. In particular, one unique regressions was estimated following our preferred specification, omitting one school-grade in each iteration. The figure suggests that the results are not driven by one particular school-grade. Indeed, the distribution is very tight and center around the value of the coefficient found in the main analysis.

Finally, we analyze whether having a higher proportion of high achievers increases the probability that a child attrits from our school sample between grades. The potential threat is that differential attrition could present an estimation challenge. Appendix Table E.4 shows the impact of the (leave-one-out) proportion of high achievers on the likelihood of leaving the sample between two consecutive grades. It shows that the children in are no more likely to attrit when they are exposed to a higher proportion of high achievers. Therefore, we do not see evidence of selective attrition. Nevertheless, in Appendix Table E.5, we restrict our

[20]We also use the parent's education to construct a proportion of high achievers similar to the used by Cools et al. (2022). Appendix E shows that the results are similar to those obtained using the main specification.

sample to the balanced panel of children and estimate the main equation. We find that the results are similar.

3.2 Dynamics

Due to the uniqueness of the data, we can study the dynamics of peer effects and how they accumulate over time. For this, we estimate the following specification:

$$y_{isc,t+l} = \beta_1 FracHighAchievers_{sct} + \mathbf{X}_{isct}\alpha + \delta_{st} + \varepsilon_{isct} \tag{3}$$

where $y_{isc,t+l}$ is one of students' i performance (measured by the corresponding test score) at the end of grade $t+1$. We estimate these regressions for every grade (Kindergarten to 6th grade). When $l=0$, equation 3 is equivalent to equation 1 and provides estimates of the short-run (contemporary) effects of peer composition at grade t, $FracHighAchievers_{sct}$, on learning at the end of that same grade, y_{isct}. We label this effect as $\beta_{t,0}$. When $l>0$, equation 3 provides estimates of the medium-term effects of peer composition at various lags (at most 6 for Kindergarten), which we label $\beta_{t,l}$

4 Results

4.1 Main Results

We estimate the regression model specified in Equation 1 above for each cognitive and non-cognitive test score. Table 3 shows that the proportion of high achiever peers reduces the scores in math and executive Function skill tests. For instance, in math, a one standard deviation increase in the leave-one-out proportion of high achievers reduces math test scores by 0.011 SD. Similarly, an increase by a standard deviation is associated with a 0.014 standard deviation decrease in the executive function test. These results are surprising given the previous literature on high achievers that found mostly positive effects but are consistent with those found by Hoxby and Weingarth (2005). Regarding the magnitude, the effects are slightly smaller than the 0.04SD decrease in GPA found by Chen and Hu (ming). However, their estimate is for high-ability students which could explain the differences as we explain later. There are two plausible explanations for this effect. First, higher achieving peer depress performance but only for those pushed to a lower rank in the local distribution, consistent with previous findings on tracking (de Roux and Riehl, 2022). Second, students are competing, which affects their mental health, motivation, and stress levels, leading to lower results, or teachers neglect other students when there are high achievers in the classroom, consistent with previous findings on the competition effects in education (Chen and Hu, ming). Finally, Table A.1 shows that the proportion of high achievers negatively affects each of the three domains of math and 2 out of 3 of the domains in executive function.

Given the asymmetric gender effects of peers found in the literature, Figure 3 shows that, on average, the coefficients for the proportion of males tend to be larger in magnitude than those for females, but the differences are not signficant. Furthermore, the figure shows how the share of females and males affects both fellow males and females separately. Interestingly, the results indicate that the proportion of male high achievers only affects male students, while the proportion of female high achievers affects only female students. These results are consistent with some of the findings of Lavy et al. (2012) that found that high-achieving male peers negatively impact boys and suggest that crowding out of top-tier activities drives the results.

These findings may seem unexpected considering that Busso and Frisancho (2021) and Cools et al. (2022) find that male students negatively affect female ones. However, it is important to note that those studies use a sample of middle and high school students. The interaction between males and females is high in middle and high school, which could explain why males impact females. This study uses a sample of elementary school children. During this period, female students tend to befriend other females, and male students tend to befriend other males (McPherson et al., 2001; Garrote et al., 2023). Indeed, in our sample, 89.6% of the males report having a male best friend and 92.7% of the females report having a female best friend. Moreover, around 19% of the children have opposite gender friends. Therefore, this behavior could explain why the effects are concentrated mostly within gender, as same-gender friend networks are more likely during these years.

Given that teachers identify high achievers without knowing their results in the test score, mean reversion is unlikely to explain our results for two reasons. Our preferred measure of high-achieving children relies on the rankings given by multiple teachers instead of just one, so no explicit design allows teachers to strategically adjust who is considered a high achiever. Second, complementary results suggest that, on average, teachers do not make mistakes when identifying high achievers, and there is no evidence suggesting that teachers across grades differ on how good they are at identifying higher-achieving students.

However, recognizing there could be a mechanical reason for why test scores could decrease, we address mean reversion by analyzing changes in the distribution of children at the top end. Over the sample period, math test scores exhibit some mean reversion: children who were in the top 20% of scores the previous year experience a 9 percentile decrease in their math scores in the current year. This is slightly higher than the 5 percentile decrease observed in Lavy et al. (2012). Nevertheless, the random assignment should have balanced out any differences in test scores. Therefore, conditional on ability and being above the 80th percentile, children should be similarly affected by mean reversion, regardless of peer composition, particularly the proportion of high achievers. Indeed, when dividing the top quintile into 20 percentiles, the decrease ranges from 7 to 10 percentiles, with most clustered around 9. Therefore, it is unlikely that statistical or incidental mean reversion explains our results.

4.2 Robustness checks and Alternative Specifications

This section provides a brief summary of a number of alternative specifications estimated and robustness checks to address several concerns about the main specification. The initial three specification checks test whether the results are robust to different ways of controlling for differences across schools and grades. The first check excludes the grade fixed effect and includes only the school fixed effect. The second check includes separately the school and grade fixed effects. Finally, the last specification check includes teacher controls to account for how teachers change the classroom environment and affect the students outcomes with their interactions. Figure 4 present the results of these alternative specifications checks. Following the main result (i.e., the main specification), the next three estimates demonstrate the insensitivity to the choice of fixed effects.

The next robustness checks test the sensitivity of the estimates to different ways of defining who is a high achiever in the classroom. First, our preferred measure uses information from all the teachers in the previous grades and assigns a child as a high achiever only if at least 50% of the teachers who observed them in previous grades identified them as such. However, given that this cutoff was chosen arbitrarily, we check how robust are the results to the use of other cutoffs. Second, we construct a proportion of high achievers using the information from the previous teacher only instead of using all the previous teachers. The next four estimates in Figure 4 show the main specification using 25%, 75% and 100% as the cutoffs, and using the previous grade teacher only, respectively. The four coefficients show that the estimates are not sensitive to the choice of cutoff.

Finally, we analyze the sensitivity of the results to constructing the proportion of high achivers using test scores. We construct two different measures: (i) whether a student is over the 95th percentile of the corresponding test score the previous year, and (ii) whether a student is over the 95th percentile of the baseline PPVT test. The last two coefficients in Figure 4. The coefficient using the previous grade test score shows similar results, while the one using the PPVT score is similar in magnitude but not significant. However, this coefficient should be interpreted cautiously, as this measure was collected only at the baseline and the new entrants do not have information. Also, the PPVT is a vocabulary test which could also explain the absence of effects on math.

4.3 Dynamics

We estimate the dynamics of the effects by estimating the regression of peer composition separately for children in each grade (Kindergarten to 6th grade), both contemporaneously (without lags, as in equation 1 above), and at various lags (at most 6 as in equation 3). Table 6 presents the results for the proportion of high achievers in math. Column (1) shows that the short-term effect of the proportion of high achievers is negative overall but only significant for 1st grade. Regarding the evolution over time of the effects, in columns (2) through (5), we report estimates of the effect of the proportion of high achievers on math achievement after (up to) 1, 2, 3, 4, and 5 lags, respectively. In the first row, corresponding to peer composition in 1st grade, the coefficients are slightly stable and significant. However, as opposed to previous evidence on ranking (Carneiro et al., ming), the effects over time decline. Interestingly, the table suggests that only 1st grade matters and that peer composition during this early grade is particularly relevant. These results are aligned with previous evidence on how Kindergarten experiences have long-run consequences (Rury, 2022). Similarly, Table 7 presents the results on Executive function, showing similar findings to math test scores for 1st grade, but also presenting effects in subsequent grades. These results suggest that exposure to high achievers has long-term effect on child Executive function, a set of skills related to a child's ability to plan, focus attention and remember instructions.

4.4 Mechanisms

Having shown that the proportion of high achievers negatively affects test scores both on cognitive (math and Executive Function), we explore potential mechanisms that might explain our results. We focus on two main channels: (i) the competitive classroom environment, (ii) the effects on other outcomes (non-cognitive skills, happiness, and effort), and (iii) the role of teacher quality.

4.4.1 Competitive Environment

To explore the role of student competition in driving our results, we first examine the heterogeneity by analyzing variations in the results across previous test scores in $t-1$. Second, we estimate the effects, taking into account the intensity of the competition given by the non-linear effects of exposure to varying proportions of high achievers. Third, we also estimate the effects separately for small and large schools to examine the varying intensity of competition. Finally, we analyze the interaction between the two sources of competition.

In our context, children competing with high achievers (i.e., children with higher ability in the previous year) are more likely to be pushed to a lower rank. Figure 5 divides the sample into five quintiles based on the previous year's math test scores. It shows that high achievers negatively affect only those students in the top quintile. This result aligns with literature suggesting that high achieving students are significantly more competitive compared to low achieving students (Syeda and Khalid, 2012). Hence, if competition is likely to play a role, then the prediction of the "Invidious Comparison Model" should hold, showing an impact of competition especially for those who were already at the top of the distribution. Also, this is consistent with previous findings that high-ability students are the ones negatively affected by high achievers given the competition among them (Chen and Hu, ming)

Second, we estimate the non-linear effects of peer composition. For this, we create quintiles based on the proportion of high achievers a child is exposed to in a given classroom. We estimate the effects of those quintiles on test scores, using the first quintile as the base level. Figure 6 shows that the effects of high achievers are negative only for those in the top three quintiles of the distribution for both math and executive function. Indeed, having fewer than three high achievers in the classroom does not have negative effects, but having more does. This suggests that competition might play a role, as children feel more pressure to perform or experience crowding-out effects due to the limited number of slots for top-tier activities. These results also suggest that the extra competition produced by an increase in the proportion of high achievers negatively impacts the test scores for cognitive skills.

Third, we estimate the effects by size of the school. For this, we calculate the average size of the school over the 7 years of the experiment and separate the schools into two groups (small and big). We expect the benefits to outperform classmates to be bigger in smaller schools. This implies that in smaller schools, children are more likely to perceive their classmates as potential competitors. Figure 7 illustrates that the detrimental effects of high achievers diminish as the competition pool increases. Indeed, the estimates between smaller and bigger schools are statistically different.

Finally, we interact the non-linear effects with the school size to examine how the peer effects change when the two sources of variation in competition interact. Intuitively, the effects should be reinforced. Panel A in Figure A.4 shows that the non-linear effects are stronger when restricting the sample to the smaller schools only. However, there are no significant effects when restricted to the bigger schools (Panel B in Figure A.4).

4.4.2 Effects on other outcomes

In the previous section, we show that the peer effects vary with the level and intensity of competition. In this section, we explore how high achievers affect other outcomes relative to non-cognitive skills and motivation. First, we estimate the impacts on non-cognitive skills to determine whether the presence or absence of competition is related to the cognitive skills results. Table 4 shows no significant effects on the non-cognitive skills measured, such as depression, self-esteem, grit, and growth mindset in 6th grade. However, these estimates should be interpreted cautiously given our smaller sample size, the estimates are noisier and our power to detect effects is smaller. One potential explanation is that

Now, we turn to self-reported measures of happiness and effort. One potential explanation of the results could be that children feel less happy when they experience more competition leading them to put in less effort in their work and tests. If students put in less effort, then the estimations of the effects on the test scores would need to account for that. Table 5 shows that self-reported happiness levels decrease as the leave-one-out proportion of high achievers increases. However, this reduction in happiness is not accompanied by a decline in self-reported effort levels. Interestingly, Appendix Figure A.5 shows that self-reported happiness only decreases for those at the top of the distribution in the previous year. Therefore, these results suggest that reductions in test scores could be attributed to decreased motivation, as indicated by lower self-reported happiness, but not to a decrease in self-reported effort.

4.4.3 Teacher Quality

Given the previous literature on teacher quality and how it increases test scores, we test whether teacher behaviors, as measured by the CLASS, mitigate the negative peer effects. Panel A in Figure 8 shows that high-quality teachers can reduce the negative effects but still have some detrimental consequences on math test scores. Panel B in the same figure examines the results by domain of the CLASS and illustrates that teachers who are high-quality on the emotional support and classroom organization are the ones that can mitigate the effects. This suggests that teachers with better-planned classroom routines and structure can lessen the negative impacts of higher achievers. These results suggest that teacher quality matters and that some neglect might occur in the classrooms, affecting the test scores in cognitive skills.

5 Conclusion

This paper analyzes the impact of peer and gender composition on students' cognitive and non-cognitive skills in elementary school settings. Data comes from a unique longitudinal experiment in Ecuador where children are randomly assigned to classrooms at the beginning of the year for 7 consecutive years, implying that the proportion of high achievers in the classroom within a school is random. Notably, in our data, two students with the same underlying ability and attending the same school can have different peer composition because they are randomly assigned to different classrooms, with slightly different peers.

We find that exposure to high achievers can negatively affect test scores in cognitive skills such as math and executive function, particularly for students who were top performers in the previous year. The negative effects are more pronounced in classrooms with a higher proportion of high achievers, suggesting that increased competition may play a significant role. However, no significant effects were observed on non-cognitive skills which suggest that future research should explore how high achievers affect non-cognitive outcomes. These findings contribute to understanding how peer and gender composition influence student outcomes during formative years. Finally, having higher-quality teachers can help mitigate the adverse effects on learning.

This research has several advantages compared to previous studies. First, we can take advantage of the random assignment in elementary school to further analyze peer composition effects in a developing country in Latin America. This setting allows us to explore the variation in the proportion of high achievers in the classroom, overcoming any selection and reflection biases in other peer effect studies. Second, we show that peer composition has long-term effects, but these decline over time, again by leveraging a nearly perfect random (98.9%) allocation of students to classrooms for seven consecutive years. Third, we explore both competition and teachers' quality as mechanisms behind the observed effects of exposure to high-achiever peers, refining insights into the complex dynamics in early educational settings.

Finally, given that peer effects are context-specific, these results raise essential policy concerns about properly addressing these adverse effects in Latin America and understanding what policies work better. Future research should explore the underlying mechanisms through which exposure to high achievers affects test scores and identify ways to mitigate this through optimal allocation of children to classrooms or by providing the teachers with the necessary tools to address them.

References

Aizer, A. (2008). Peer Effects and Human Capital Accumulation: The Externalities of ADD. *NBER Working Paper No. 14354*.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., and Schady, N. (2016). Teacher Quality and Learning Outcomes in Kindergarten. *Quarterly Journal of Economics*, 131(3):1415–1453.

Araujo, M. C., Dormal, M., and Schady, N. (2019). Childcare Quality and Child Development. *Journal of Human Resources*, 54(3):656–682.

Balestra, S., Sallin, A., and Wolter, S. C. (2023). High-Ability Influencers? The Heterogeneous Effects of Gifted Classmates. *Journal of Human Resources*, 58(2).

Bassi, M., Meghir, C., and Reynoso, A. (2020). Education Quality and Teaching Practices. *Economic Journal*, 130(631):1937–1965.

Berlinski, S. and Schady, N. (2015). *The Early Years: Child Well-Being and the Role of Public Policy*. Palgrave Macmillan, New York, NY.

Bertoni, M., Brunello, G., and Cappellari, L. (2020). Who Benefits from Privileged Peers? Evidence from Siblings in Schools. *Journal of Applied Econometrics*, 35(7):893–916.

Bietenbeck, J. (2020). The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR. *Journal of the European Economic Association*, 18(1):392–426.

Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit Theories of Intelligence Predict Achievement across an Adolescent Transition: A Longitudinal Study and an Intervention. *Child Development*, 78(1):246–263.

Busso, M. and Frisancho, V. (2021). Good Peers have Asymmetric Gendered Effects on Female Educational Outcomes: Experimental Evidence from Mexico. *Journal of Economic Behavior and Organization*, 189.

Card, D. and Giuliano, L. (2013). Peer Effects and Multiple Equilibria in the Risky Behavior of Friends. *Review of Economics and Statistics*, 95(4):1130–1149.

Carneiro, P., Aguayo, Y. C., Salvati, F., and Schady, N. (Forthcoming). The Effect of Classroom Rank on Learning throughout Elementary School: Experimental Evidence from Ecuador. *Journal of Labor Economics*.

Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). Does your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, 27(3):439–464.

Carrelll, S. E., Bruce I, S., and West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81(3):855–882.

Chen, S. and Hu, Z. (Forthcoming). How Competition Shapes Peer Effects: Evidence from a University in China. *Tbe Review of Economics and Statistics*, (3):883–931.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Edded Estimates. *American Economic Review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9):2633–2679.

Cools, A., Fernández, R., and Patacchini, E. (2022). The Asymmetric Gender Effects of High Flyers. *Labour Economics*, 79.

de Gendre, A. and Salamanca, N. (2020). On the Mechanisms of Ability Peer Effects. *IZA DP No. 13938*.

de Roux, N. and Riehl, E. (2022). Do College Students Benefit from Placement into Higher-achieving Classes? . *Tbe Review of Economics and Statistics*, (210).

Drago, R. and Garvey, G. T. (1998). Incentives for Helping on the Job: Theory and Evidence. *Journal of Labor Economics*, 16(1):1–25.

Duckworth, A. L. and Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91(2):166–174.

Dunn, L., Lugo, D., Padilla, E., and Dunn, L. (2015). *Test de Vocabulario en Imágenes Peabody*. American Guidance Service, Circle Pines, MN.

Feld, J. and Zölitz, U. (2017). Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. *Journal of Labor Economics*, 35(2):387–428.

Feld, J. and Zölitz, U. (2022). The Effect of Higher-Achieving Peers on Major Choices and Labor Market Outcomes. *Journal of Economic Behavior and Organization*, 196:200–219.

Garrote, A., Zurbriggen, C. L., and Schwab, S. (2023). Friendship Networks in Inclusive Elementary Classrooms: Changes and Stability related to Students' Gender and Self-perceived Social Inclusion. *Social Psychology of Education*.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, 118(3):1049–1074.

Hanushek, E. A., Kain, J. F., Markman, J. M., and Rivkin, S. G. (2003). Does Peer Ability affect Student Achievement? *Journal of Applied Econometrics*, 18(5):527–544.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, 103(6):2052–2086.

Heckman, J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Hoxby, C. (2000a). Peer Effects in the Classroom: Learning from Gender and Race Variation. *NBER Working Paper No. 7867*.

Hoxby, C. (2000b). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of Economics*, 115(4):1067–1101.

Hoxby, C. and Weingarth, G. (2005). Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects. *Unpublished Manuscript, Harvard University*.

Jacob, B. A., Lefgren, L., and Sims, D. P. (2010). The Persistence of Teacher-induced Learning. *Journal of Human Resources*, 45(4):915–943.

Jochmans, K. (2023). Testing Random Assignment to Peer Groups. *Journal of Applied Econometrics*, 38(3):321–333.

Lavy, V., Silva, O., and Weinhardt, F. (2012). The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics*, 30(2):367–414.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27:415–444.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H. L., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., and Caspi, A. (2011). A Gradient of Childhood Self-control predicts Health, Wealth, and Public Safety. *Proceedings of the National Academy of Sciences of the United States of America*, 108(7):2693–2698.

Mouganie, P. and Wang, Y. (2020). High-Performing Peers and Female STEM Choices in School. *Journal of Labor Economics*, 38(3):805–841.

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

Näslund-Hadley, E. and Bando, R. (2015). *All Children Count: Early Mathematics and Science Education in Latin America and the Caribbean: Overview Report*. Inter-American Development Bank, Washington D.C.

Pianta, R., Paro, K. L., and Hamre, B. (2015). *Classroom Assessment Scoring System—CLASS*. Brookes, Baltimore, MD.

Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics*, 125(1):175–214.

Rury, D. (2022). Putting the K in Rank: How Kindergarten Classrooms Impact Short and Long-Run Outcomes. *SSRN Electronic Journal*.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116(2):681–704.

Sacerdote, B. (2011). Peer Effects in Education: How Might they Work, How Big are they and How much do we Know Thus Far? In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier.

Sacerdote, B. (2014). Experimental and Quasi-experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, 6:253–272.

Siegler, R. S. and Booth, J. L. (2004). Development of Numerical Estimation in Young Children. *Child Development*, 75(2):428–444.

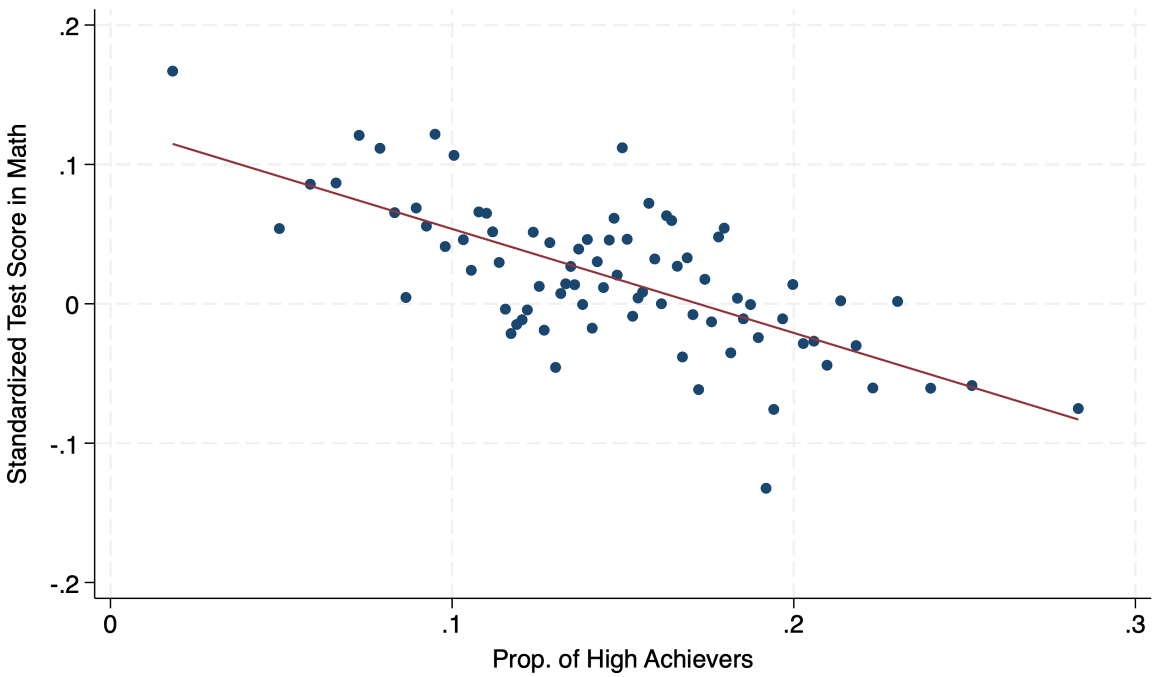
Syeda, S. H. and Khalid, R. (2012). Differences in Achievement Motivation and its Salient Components among High and Low Achieving Students. *Pakistan Journal of Psychology*, 43(1).

Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M. C., Weiland, C., Gomez, C. J., Moreno, L., Rolla, A., D'Sa, N., and Arbour, M. C. (2015). Experimental Impacts of a Teacher Professional Development Program in Chile on Preschool Classroom Quality and Child Outcomes. *Developmental Psychology*, 51(3):309–322.

6 Figures and Tables

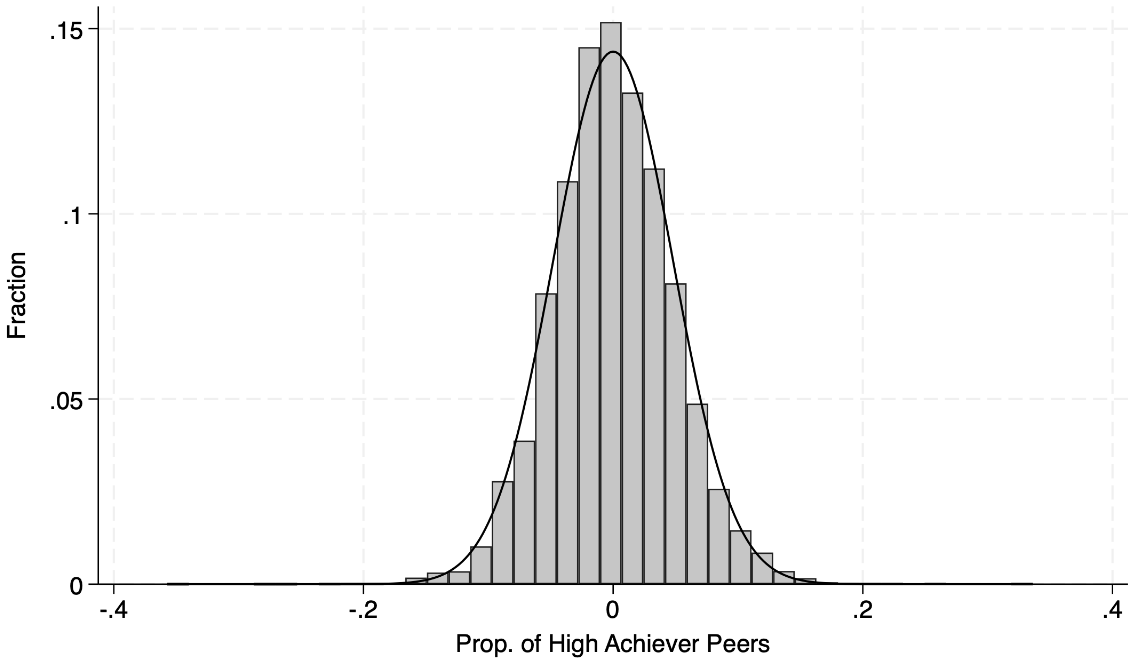
6.1 Figures

Figure 1: Correlation between Proportion of High Achievers and Cognitive Skills



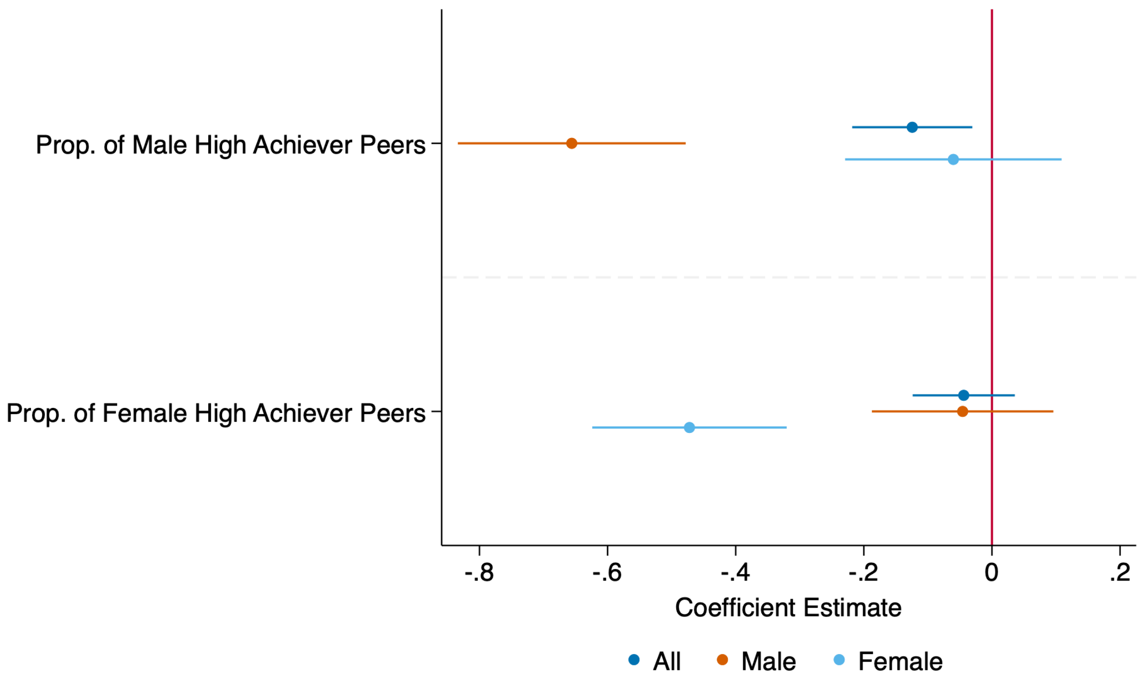
Notes: The figure above shows the relationship between the (leave-one-out) proportion of high achievers and math test scores. The plot does not include controls.

Figure 2: Residual Share of High Achievers Peers Across School-Grades



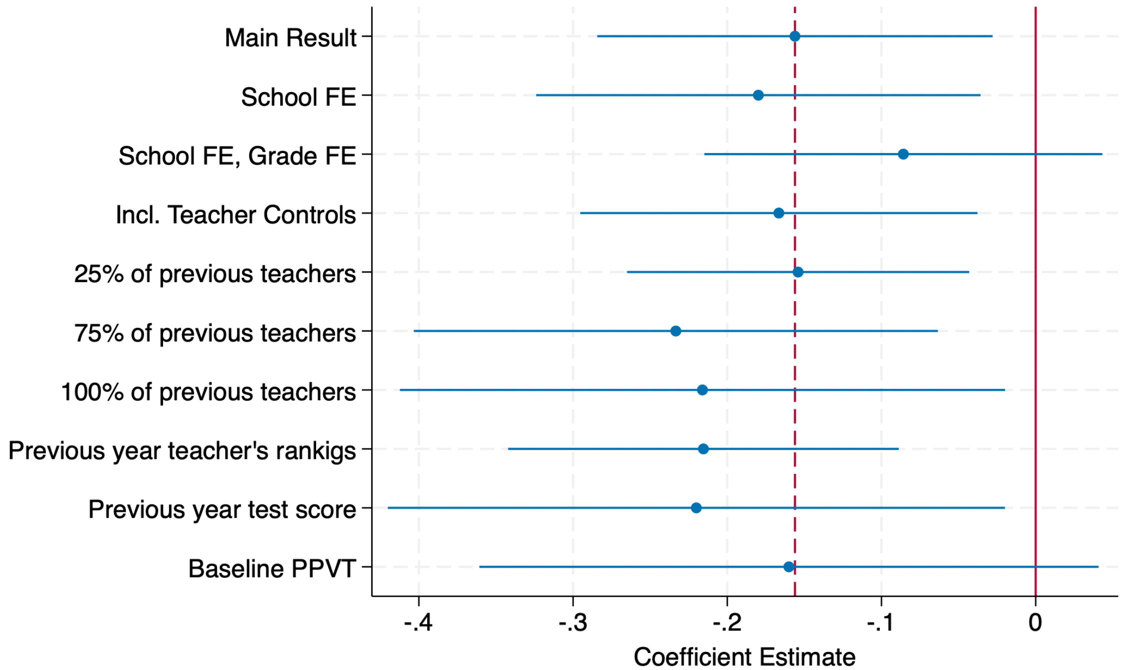
Notes: The figure above represents the residualized high achiever peer share distribution across school-grades, conditional on school-by-grade fixed effects. All regressions are limited to schools with at least two classrooms per grade. The overlaid curve represents the normal distribution.

Figure 3: Heterogeneous Effects of High Achievers on Cognitive Skills by Gender



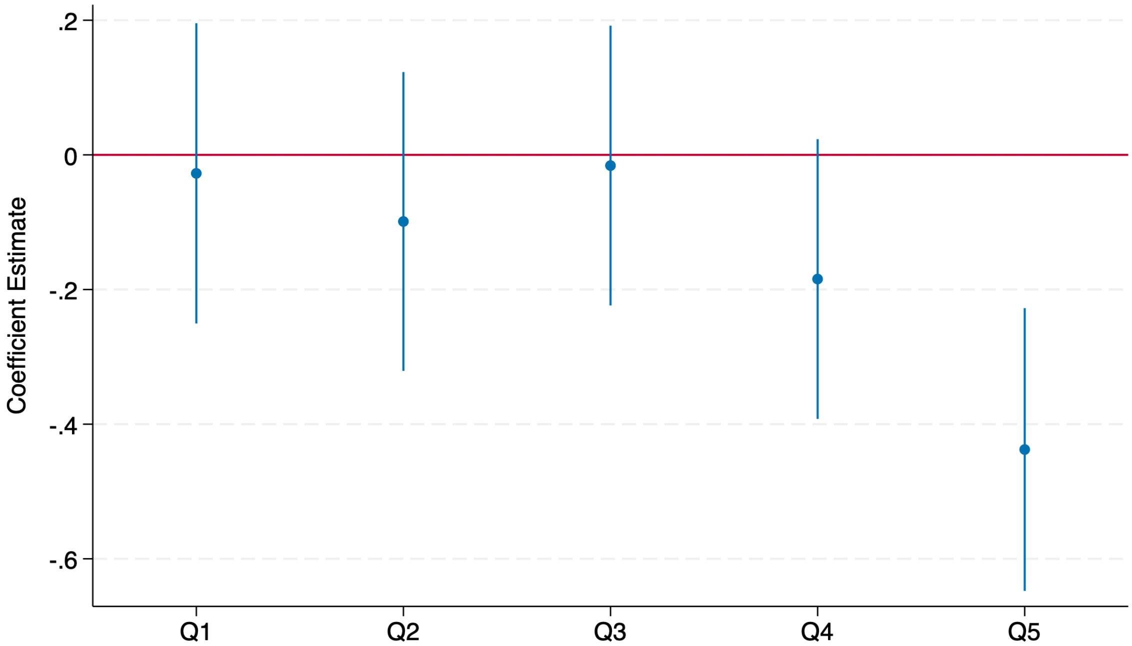
Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers separated by gender on math test scores separately for males and females. All regressions are limited to schools with at least two classrooms per grade. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Figure 4: Robustness Checks on Cognitive Skills



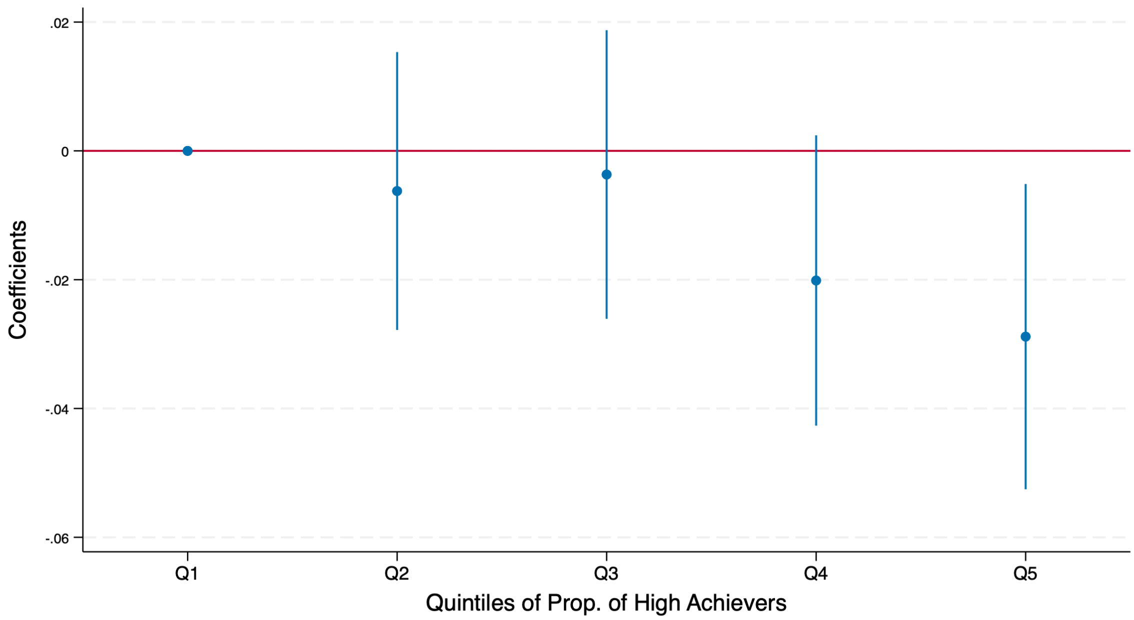
Notes: The figure above reports estimates from regressions of the leave-one-out proportion of high-achiever peers on math test scores. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school-by-grade level. The "Main Result" estimate (at the top) uses the specification Table 3. All other estimates are variations on the baseline model. Estimates 2-4 vary the set of fixed effects and controls included. Estimates 5-7 use different cutoffs to define who is a high achiever using the teachers' rankings from teachers in the previous grades. Estimate 8 uses only the information from the previous teacher to define who is a high achiever. Estimate 9 uses the previous year test scores to define who is a high achiever, while estimate 10 uses the Baseline PPVT score to define who is a high achiever.

Figure 5: Heterogeneous Effects of High Achievers on Cognitive Skills by Previous Test Results



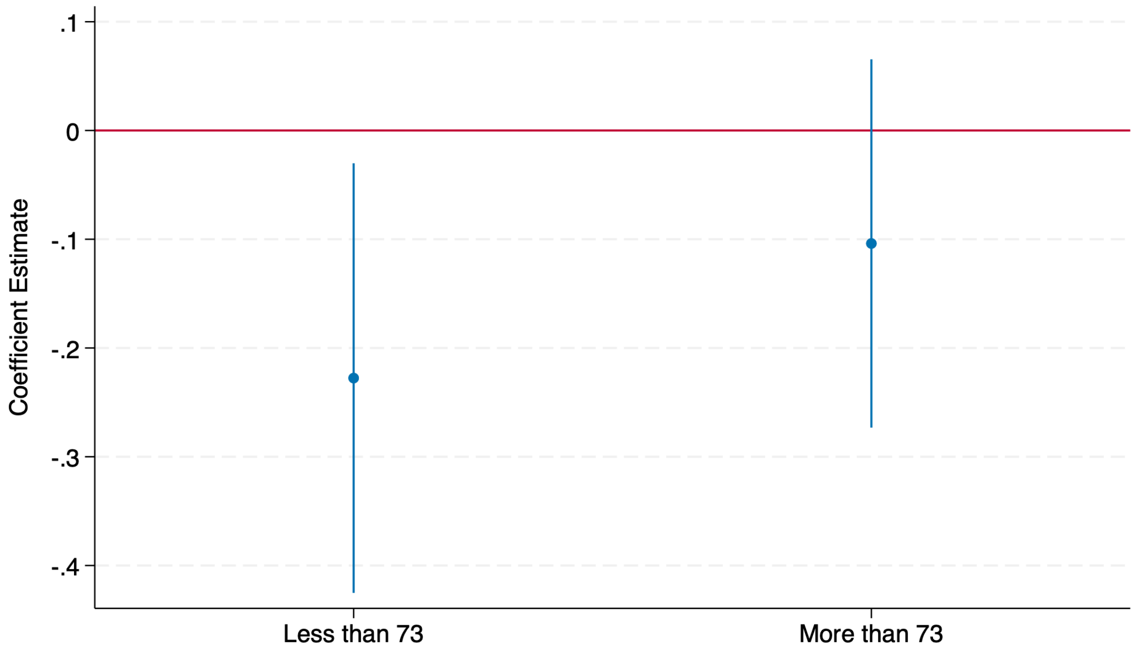
Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills by quintiles of the test result in the previous year. All regressions are limited to schools with at least two classrooms per grade. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Figure 6: Non-Linear Effects of High Achievers



Notes: The figures above report non-linear estimates from regressions of the quintiles of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

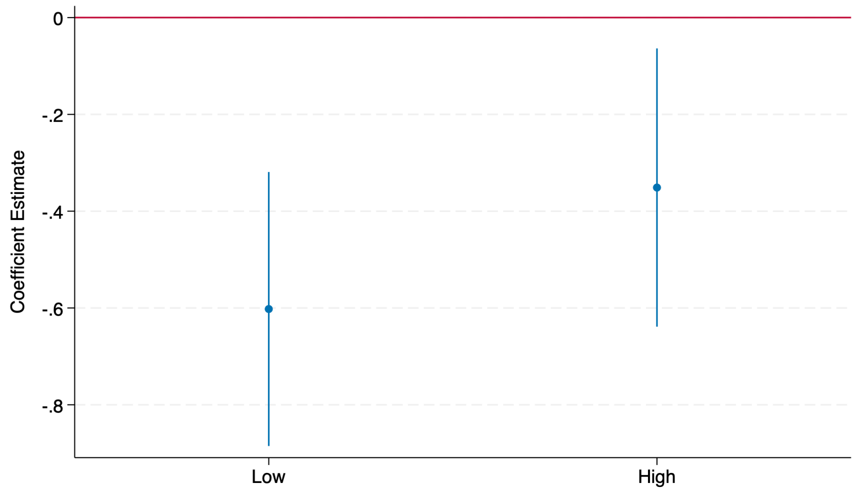
Figure 7: Heterogeneous Effects of High Achievers on by School Size



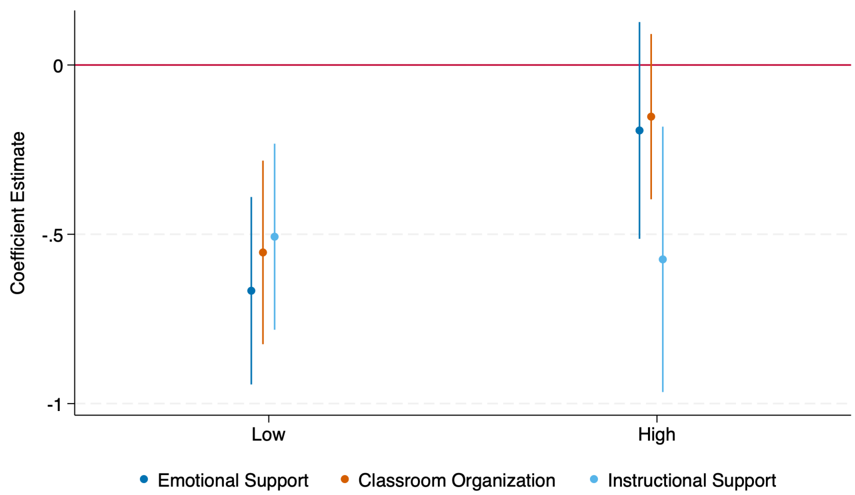
Notes: The figures above report non-linear estimates from regressions of the quintiles of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Figure 8: Heterogeneous Effects of High Achievers on Cognitive Skills by Teacher Quality

(a) CLASS Score



(b) Components



Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills by teacher quality measured using the CLASS *(Classroom Assessment Scoring System)* score. All regressions are limited to schools with at least two classrooms per grade. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

6.2 Tables

Table 1: Child, Teacher, and Classroom Characteristics

	Mean	SD	Median	Min	Max
A. Child characteristics					
Age of child (in months) in 2012	60.30	4.94	60.00	32.00	142.00
Sex (1 = Female)	0.49	0.50	0.00	0.00	1.00
Receptive vocabulary score (PPVT)	82.89	15.87	81.00	55.00	145.00
High Achiever	0.13	0.34	0.00	0.00	1.00
Lagged Math Test Score	0.00	1.00	0.02	-3.83	3.58
Lagged Executive Function Test	0.00	1.00	0.05	-5.45	6.13
Self-reported Happiness	0.90	0.30	1.00	0.00	1.00
Self-reported Effort	0.85	0.35	1.00	0.00	1.00
Proportion who attended preschool	0.60	0.49	1.00	0.00	1.00
B. Household characteristics					
Mother's years of completed schooling	8.77	3.80	9.00	0.00	22.00
Father's years of completed schooling	8.50	3.83	8.00	0.00	22.00
Mother's age	30.22	6.57	29.00	5.00	93.00
Father's age	34.54	7.89	33.00	5.00	99.00
Household has piped water in home	0.83	0.38	1.00	0.00	1.00
Household has flush toilet in home	0.46	0.50	0.00	0.00	1.00
C. Classroom characteristics					
Class size	36.94	6.47	37.00	8.00	60.00
Prop. of High Achievers	0.15	0.07	0.14	0.00	0.57
D. Teacher characteristics					
Female	0.84	0.36	1.00	0.00	1.00
Years of experience	17.63	10.35	15.58	0.08	57.00
Prop. tenured	0.80	0.40	1.00	0.00	1.00
CLASS Score	-0.02	1.00	0.02	-4.29	4.28
Emotional Support Score	-0.02	0.99	-0.05	-4.98	4.89
Classroom Organization Score	-0.02	1.01	0.06	-5.36	2.58
Instructional Support Score	-0.02	0.98	-0.29	-1.03	7.10

Notes: Table reports summary statistics of the children in the sample. It includes children's characteristics and those of their assigned classrooms and teachers.

Table 2: Characteristics of high-achieving students

	Not High Achiever	High Achiever	Difference (1)-(2)
	(1)	(2)	(3)
A. Children characteristics			
Age of child (in months) in 2012	60.157	61.153	-0.996***
	(4.953)	(4.796)	(0.118)
Sex (1 = Female)	0.475	0.562	-0.087***
	(0.499)	(0.496)	(0.012)
Receptive vocabulary score (PPVT)	81.467	91.606	-10.140***
	(15.265)	(16.694)	(0.407)
Lagged Math Test Score	-0.154	0.922	-1.076***
	(0.943)	(0.816)	(0.008)
Lagged Executive Function Test	-0.108	0.633	-0.741***
	(0.977)	(0.892)	(0.009)
Self-reported Happiness	0.895	0.950	-0.054***
	(0.306)	(0.218)	(0.006)
Self-reported Effort	0.851	0.870	-0.018**
	(0.356)	(0.336)	(0.009)
Proportion who attended preschool	0.604	0.648	-0.043***
	(0.489)	(0.478)	(0.012)
B. Household Characteristics			
Mother's years of completed schooling	8.601	9.835	-1.234***
	(3.748)	(3.950)	(0.097)
Father's years of completed schooling	8.328	9.473	-1.145***
	(3.775)	(4.007)	(0.109)
Mother's age	30.144	30.656	-0.512***
	(6.560)	(6.618)	(0.163)
Father's age	34.495	34.748	-0.253
	(7.892)	(7.853)	(0.216)
Proportion who attended preschool	0.604	0.648	-0.043***
	(0.489)	(0.478)	(0.012)
Household has piped water in home	0.828	0.841	-0.012
	(0.377)	(0.366)	(0.009)
Household has flush toilet in home	0.455	0.473	-0.018
	(0.498)	(0.499)	(0.012)

Notes: Table reports summary statistics of the children in the sample. It includes children's characteristics and those of their assigned classrooms and teachers.

Table 3: Effects of High Achievers on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of High Achiever Peers	-0.156**	-0.194**
	(0.065)	(0.081)
Mean of Dependent Variable	0.019	0.019
Treatment Effect of 1SD increase	-0.011	-0.014
Observations	87303	56626
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table 4: Effects of High Achievers on Non-Cognitive Skills

	Depression	Self-esteem	Growth Mindset	Grit
	(1)	(2)	(3)	(4)
Prop. of High Achiever Peers	-0.285	-0.057	-0.049	-0.136
	(0.258)	(0.254)	(0.262)	(0.236)
Mean of Dependent Variable	0.001	0.001	0.001	0.001
Treatment Effect of 1SD increase	-0.017	-0.003	-0.003	-0.008
Observations	7763	7763	7763	7763
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers on non-cognitive skills. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table 5: Effects of High Achievers on Happiness and Effort

	Happiness	Effort
	(1)	(2)
Prop. of High Achiever Peers	-0.098**	0.019
	(0.048)	(0.058)
Mean of Dependent Variable	0.903	0.854
Treatment Effect of 1SD increase	-0.008	0.002
Observations	12034	12034
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table 6: Cumulative Effects of High Achievers on Math Scores

	t	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$
	(1)	(2)	(3)	(4)	(5)	(6)
1st grade	-0.521**	-0.172	-0.227**	-0.128	-0.155*	-0.150*
	(0.205)	(0.120)	(0.112)	(0.107)	(0.0911)	(0.0851)
2nd grade	-0.077	-0.121	-0.133	-0.130*	-0.0664	
	(0.160)	(0.123)	(0.0870)	(0.0713)	(0.0772)	
3rd grade	-0.172	-0.317***	-0.0821	-0.107		
	(0.158)	(0.115)	(0.0814)	(0.0846)		
4th grade	-0.089	-0.140	-0.124			
	(0.154)	(0.0972)	(0.0850)			
5th grade	-0.079	0.0188				
	(0.125)	(0.0891)				
6th grade	-0.049					
	(0.113)					
Controls	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The figure reports estimates from regressions of the leave-one-out proportion of high-achiever peers in each grade. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table 7: Cumulative Effects of High Achievers on Executive Function

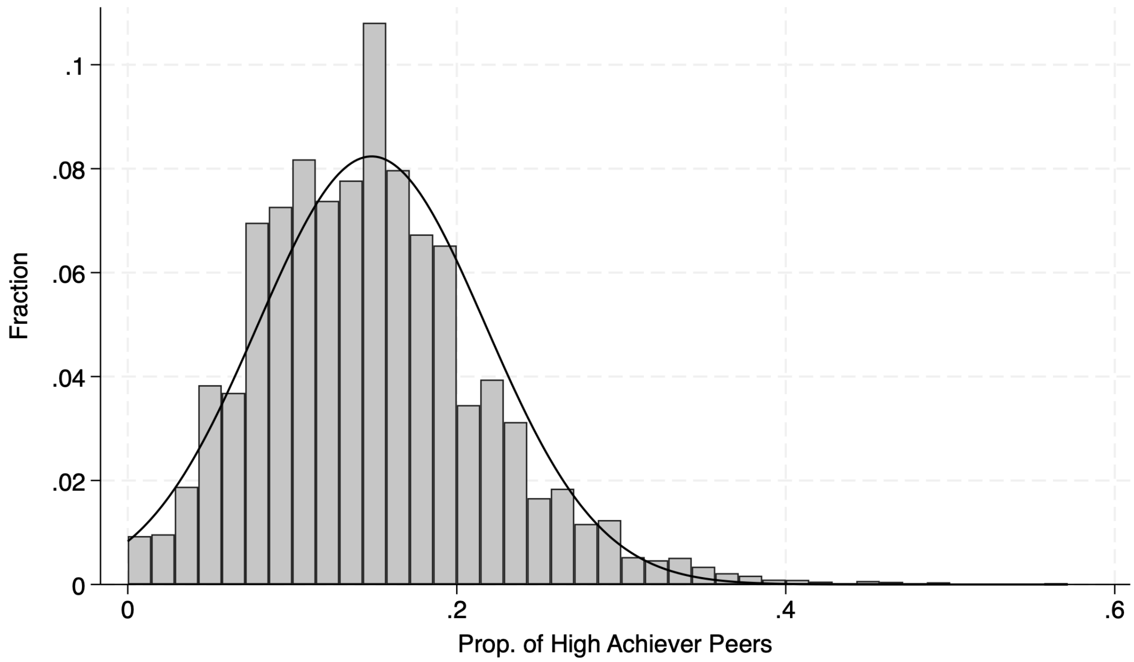
	t	$t+1$	$t+2$	$t+3$
	(1)	(2)	(3)	(4)
1st grade	-0.270*	-0.565***	-0.167	-0.513***
	(0.160)	(0.176)	(0.154)	(0.152)
2nd grade	-0.249*	-0.170	-0.289**	
	(0.146)	(0.173)	(0.121)	
3rd grade	0.160	-0.529***		
	(0.168)	(0.165)		
4th grade	-0.419**			
	(0.168)			
Controls	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes

Notes: The figure reports estimates from regressions of the leave-one-out proportion of high-achiever peers in each grade on executive function test scores. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

APPENDIX

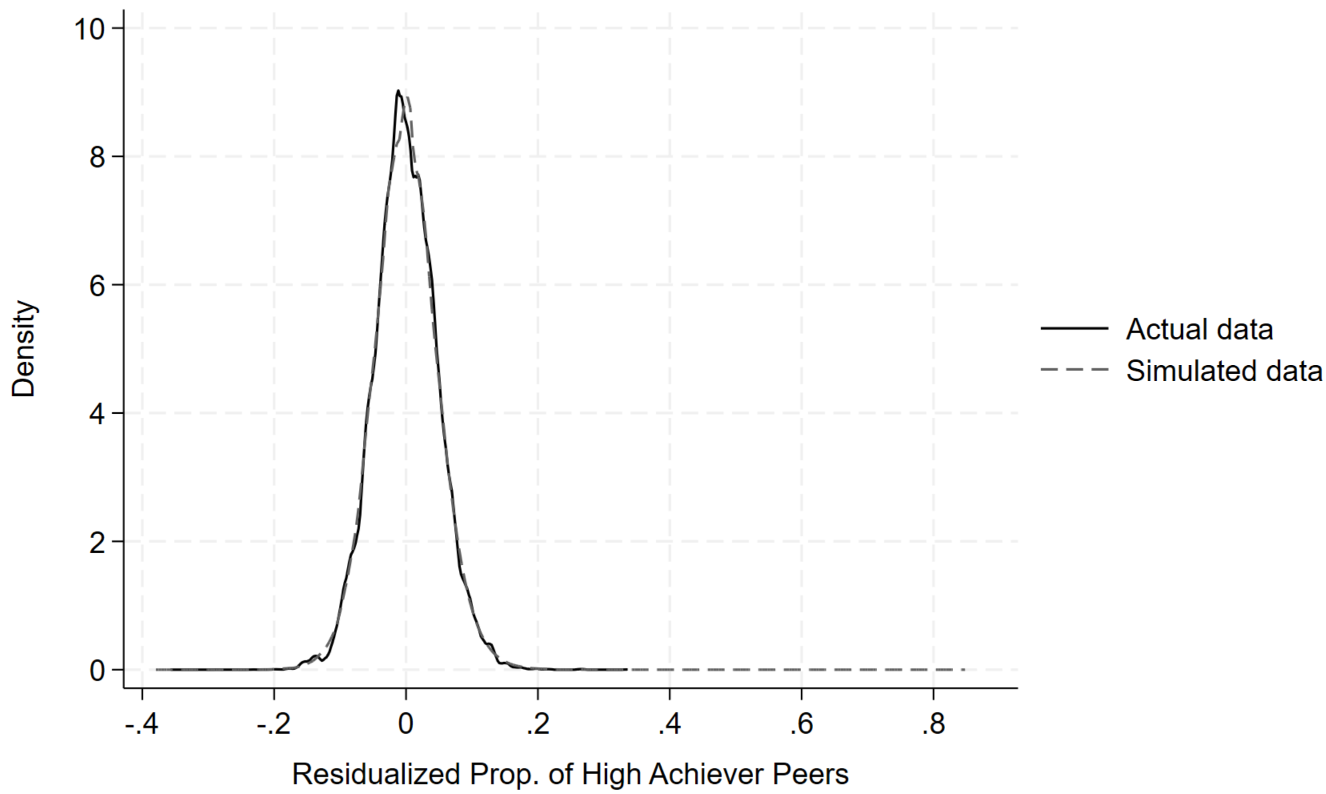
A Supplemental Figures and Tables

Figure A.1: Distribution of the Proportion of High Achievers Peers



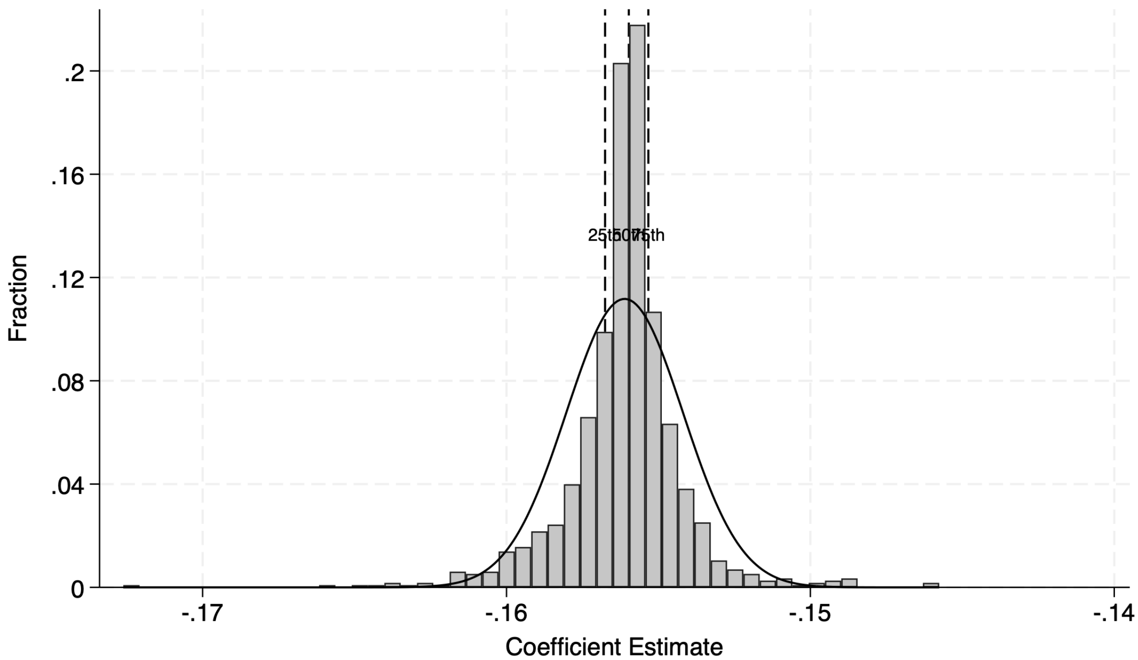
Notes: The figure above high achiever peer share distribution across school-grades. The overlaid curve represents the normal distribution.

Figure A.2: Simulated and Actual Residual Share of High Achievers Peers Across School-Grades



Notes: The figure above represents the residualized high achiever peer share distribution across school-grades, conditional on school-by-grade fixed effects. All regressions are limited to schools with at least two classrooms per grade. The overlaid curve represents the normal distribution.

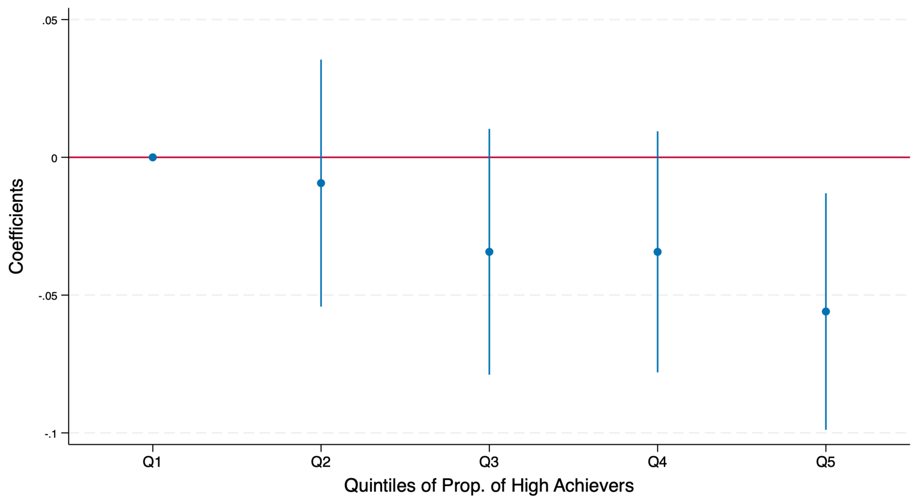
Figure A.3: Distribution of Coefficients of the Effect of High Achievers on Math



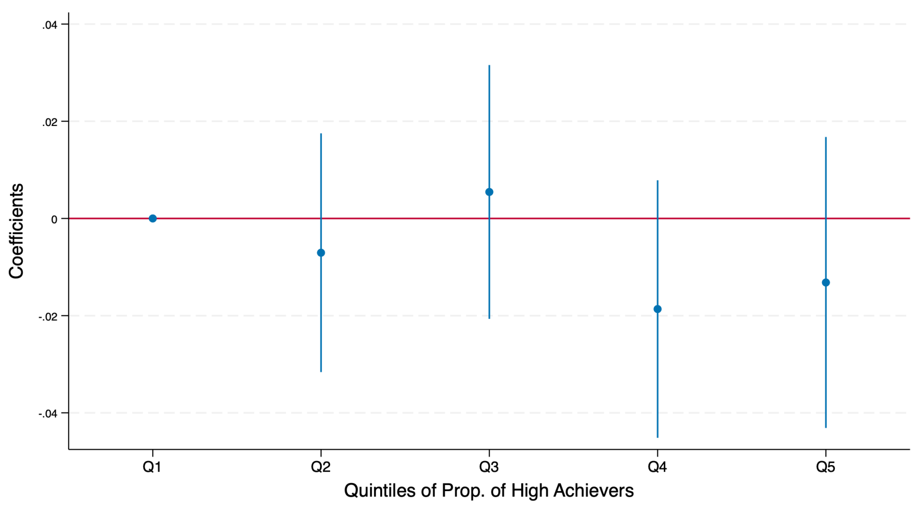
Notes: The figure above represents the distribution of the coefficients of leave-one-out regressions using the main specification. Each line represents the main specification with one school-grade committed from the sample. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. The overlaid curve represents the normal distribution.

Figure A.4: Interaction of Non-Linear Effects and School Size

(a) Small Schools



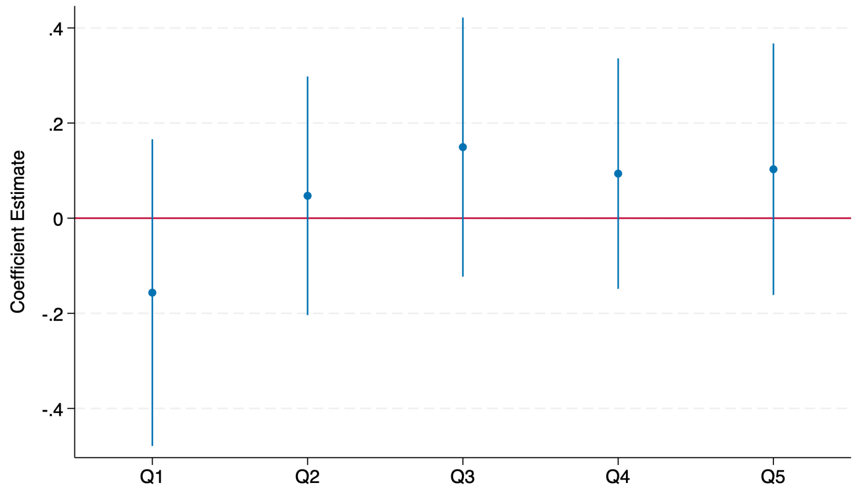
(b) Big Schools



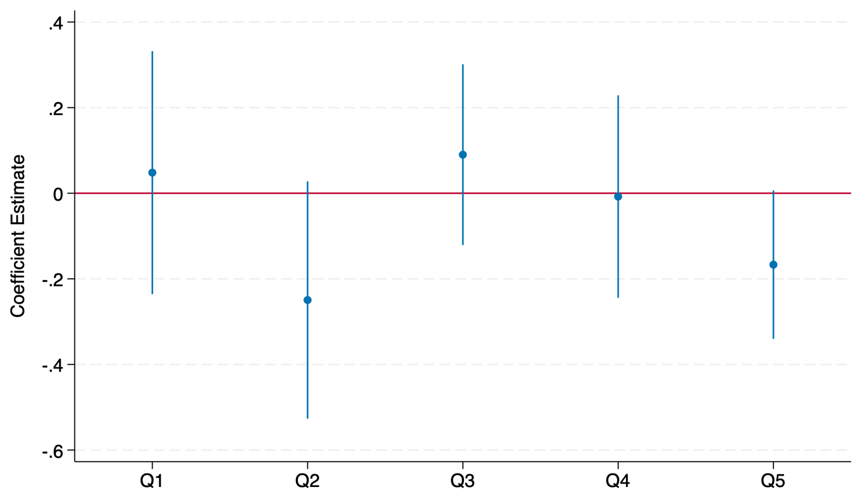
Notes: The figure above represent the distribution of the coefficients of separate regressions using the main specification on each school-by-grade. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. The overlaid curve represents the normal distribution.

Figure A.5: Effects by Prev

(a) Effort



(b) Happiness



Notes: The figure above represent the distribution of the coefficients of separate regressions using the main specification on each school-by-grade. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. The overlaid curve represents the normal distribution.

Table A.1: Effects on Cognitive Test Scores Components

	Math			Executive Function		
	Number recognition and arithmetic	Number sense	Word problems	Cognitive Flexibility	Working Memory	Inhibitory Control
	(1)	(2)	(3)	(4)	(5)	(6)
Prop. of High Achiever Peers	-0.165*	-0.255***	-0.254***	-0.235***	-0.182**	-0.186
	(0.095)	(0.074)	(0.068)	(0.087)	(0.083)	(0.123)
Mean of Dependent Variable	0.017	0.020	0.013	0.011	0.016	0.012
Treatment Effect of 1SD increase	-0.011	-0.018	-0.018	-0.017	-0.013	-0.011
Observations	87225	87093	87303	56626	56626	29984
Controls	Yes	Yes	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of peers based on their parents' education. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

B Randomization details

Table B.1: Testing for random assignment of children to classrooms, math

	Kindergarten	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6
Test statistic	1.359	-0.383	0.905	0.300	-0.445	-0.222	0.980
P-value	0.174	0.702	0.366	0.764	0.657	0.825	0.327

Notes: The table reports results for tests of random assignment of children to classrooms within schools using a methodology proposed by Jochmans (2023). The null hypothesis is the absence of correlation between a child's ability measured at the end of the previous grade and the average ability of classroom peers assigned to her at the beginning of a given grade, conditional on school. The sample includes all children.

C Details on the Tests and IRT

In this appendix, we cover more details on tests applied and the IRT procedure implemented to calculate the math test scores. First, Table C.1 and Table C.2 show summary statistics of each of the tests applied by grade for math and executive function, respectively.

We normalize the end-of-year tests by subtracting the mean and dividing by the national sample's standard deviation. We then create three test aggregates for math and executive function, respectively. Each of the four tests within an aggregate receives the same weight. Like the underlying tests, the aggregates are normalized to have zero mean and unit standard deviation.

Table C.1: Summary Statistics for Math Test Score Components

	N	Mean	SD
A. Kindergarten			
Number identification	14522	0.372	0.229
Block rotation	14522	0.805	0.156
Sequences	14522	0.256	0.299
Word problems	14512	0.252	0.163
B. 1st Grade			
Number identification	16158	0.653	0.175
Arithmetic	17265	0.362	0.246
Word problems	16353	0.491	0.254
Number line	16368	0.782	0.123
C. 2nd Grade			
Sequences	18481	0.399	0.234
Position Value	16846	0.366	0.136
Arithmetic	18481	0.406	0.230
Word problems	18393	0.300	0.190
Number line	16874	0.841	0.073
D. 3rd Grade			
Sequences	17521	0.496	0.231
Word problems	17521	0.390	0.208
Position Value	17521	0.386	0.182
Arithmetic	17521	0.527	0.227
Number line	17277	0.847	0.095
E. 4th Grade			
Sequences	17432	0.091	0.074
Word problems	17424	0.078	0.101
Position Value	17424	0.058	0.067
Arithmetic	17426	0.205	0.150
F. 5th Grade			
Sequences	17529	0.538	0.219
Word problems	17529	0.481	0.208
Position Value	17529	0.453	0.201
Arithmetic	17529	0.455	0.220
G. 6th Grade			
Sequences	17266	0.483	0.240
Word problems	17266	0.411	0.178
Position Value	17266	0.564	0.251
Arithmetic	17266	0.466	0.209

Notes: The table presents the results from pairwise correlations between CLASS score and its components collected from Kindergarten to 4th grade. All the correlations are significant at the 1 percent level.

Table C.2: Summary Statistics for Executive Function Score Components

	N	Mean	SD
A. Kindergarten			
Memory	14522	0.269	0.208
Card sorting	14511	0.799	0.225
Day and night	14506	0.848	0.242
Indicators comprehension	14519	0.617	0.166
B. 1st Grade			
Memory	17227	0.310	0.186
Card sorting	17227	0.857	0.247
Pair Cancellation	16347	0.240	0.095
Matrix	17227	0.485	0.294
C. 2nd Grade			
Memory	16839	0.445	0.197
Card sorting	18393	0.696	0.232
Pair Cancellation	16848	0.370	0.114
Words and colors - Stroop	14354	0.218	0.071
Numbers and amounts - Stroop	18393	0.833	0.270
D. 3rd Grade			
Triangles and squares	17518	0.695	0.163
Memory	17518	0.425	0.202
Pair Cancellation	17279	0.428	0.109
Words and colors - Stroop	16142	0.240	0.061
E. 4th Grade			
Triangles and squares	17424	0.768	0.146
Memory	17425	0.466	0.161
Pair Cancellation	17434	0.454	0.112
Words and colors - Stroop	16920	0.271	0.066

Notes: The table presents the results from pairwise correlations between CLASS score and its components collected from Kindergarten to 4th grade. All the correlations are significant at the 1 percent level.

D Application of the CLASS in Ecuador

CLASS Protocol

In this appendix, we cover minute details of the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2015) application protocol and how it was applied in Ecuador. We use the CLASS to measure teacher behaviors. The CLASS measures teacher behaviors in three domains: Emotional Support, Classroom Organization, and Instructional Support. Emotional support includes children's emotional and social expressions in the classroom, Classroom Organization relates to the classroom routines and teachers' proactiveness, and Instructional Support is related to promoting order thinking and providing quality feedback.

Each of the domains is composed of three of four dimensions that are scored separately by the coders. Figure D.1 provides an overview of the dimensions included in each domain. For each dimension, the CLASS protocol gives coders concrete guidance on whether the score given should be "low" (scores of 1–2), "medium" (scores of 3–5), or "high" (scores of 6–7). Each domain is scored individually, and the coders look for specific behaviors. For example, within the behavior management dimension, a coder would assess whether there are clear behavior rules and expectations and whether these are applied consistently. The CLASS protocol would assign a high score to a teacher whose rules and expectations for behavior are clear and consistently enforced. In contrast, a teacher whose rules and expectations are absent, unclear, or inconsistently enforced would be assigned a low score. (See Appendix Table B1 in Araujo et al. (2016) for more details). Under the CLASS protocol, the scoring process does not consist of running down a checklist of the presence or absence of certain behaviors or indicators but a holistic and composite description of the classroom experience.

Application in Ecuador

We filmed all teachers for an entire school day (from approximately eight to one in the afternoon for morning schools and from two to six in the afternoon for the afternoon schools).

Teachers only knew on what day they would be filmed until the day itself. Following CLASS protocols, we discarded the first hour of filming, times that were not instructional (e.g., lunch and breaks), and where the main teacher was not in the classroom (e.g., PE Class). The remaining video was cut into usable 20-minute segments in which at least five children and their main teacher were in the segment for at least 15 of the 20 minutes that the segment ran. For each teacher, we selected the first four segments that comply with the protocol (over 9,000 segments in total).

Once the videos were separated into segments, a group of 6-8 coders explicitly trained for this purpose by a Teachstone-certified CLASS K-3 trainer scored them. The trainer also provided feedback and supervised the coders while coding the segments. All the segments were double-coded by two independent coders who scored each segment on the ten dimensions explained previously. The videos with large differences in their scores between the two coders were flagged and sent to a third coding process by an independent coder.[21] Additionally, during the entire process, we interacted extensively with the developers of the CLASS at the University of Virginia.

Figure D.2 graphs univariate densities of the distribution of CLASS score and each domain by grade. The figure shows that teachers score the highest in Classroom Organization, with teachers distributed in the "medium" and "high" parts of the distribution; somewhat lower in Emotional Support, with most teachers in the "medium" range; and lowest in Instructional Support, where all teachers have "low" CLASS scores. As a result, most teachers have a medium total Score in the CLASS, and some are in the Low score range.

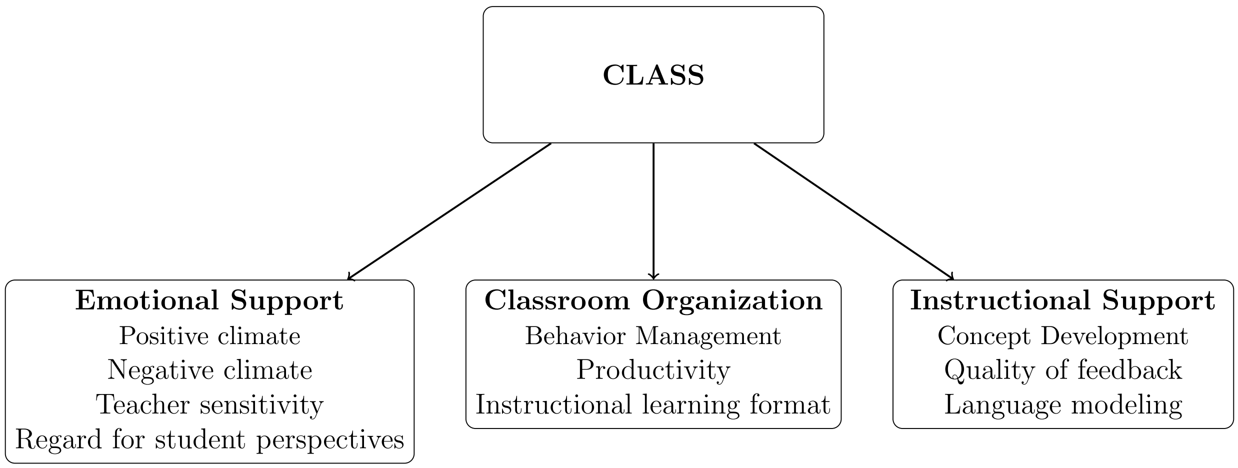
Table D.1 shows the correlation between the total CLASS score and each of its components. It shows that the components' scores are highly correlated with the total, which is consistent with how the total score is calculated. Moreover, it shows that the correlation among components is lower, consistent with the fact that each measures different behaviors

[21] Based on a preliminary analysis of CLASS data from Ecuador that revealed a breakdown of high-and low-variability dimensions, videos are coded a third time if they have a difference of more than 1 point in the dimensions of Negative Climate, Concept Development, Quality of Feedback, or Language Modeling; if there is a difference of more than 2 points in any of the other dimensions they are also flagged for re-coding

and dimensions of the teacher's quality.

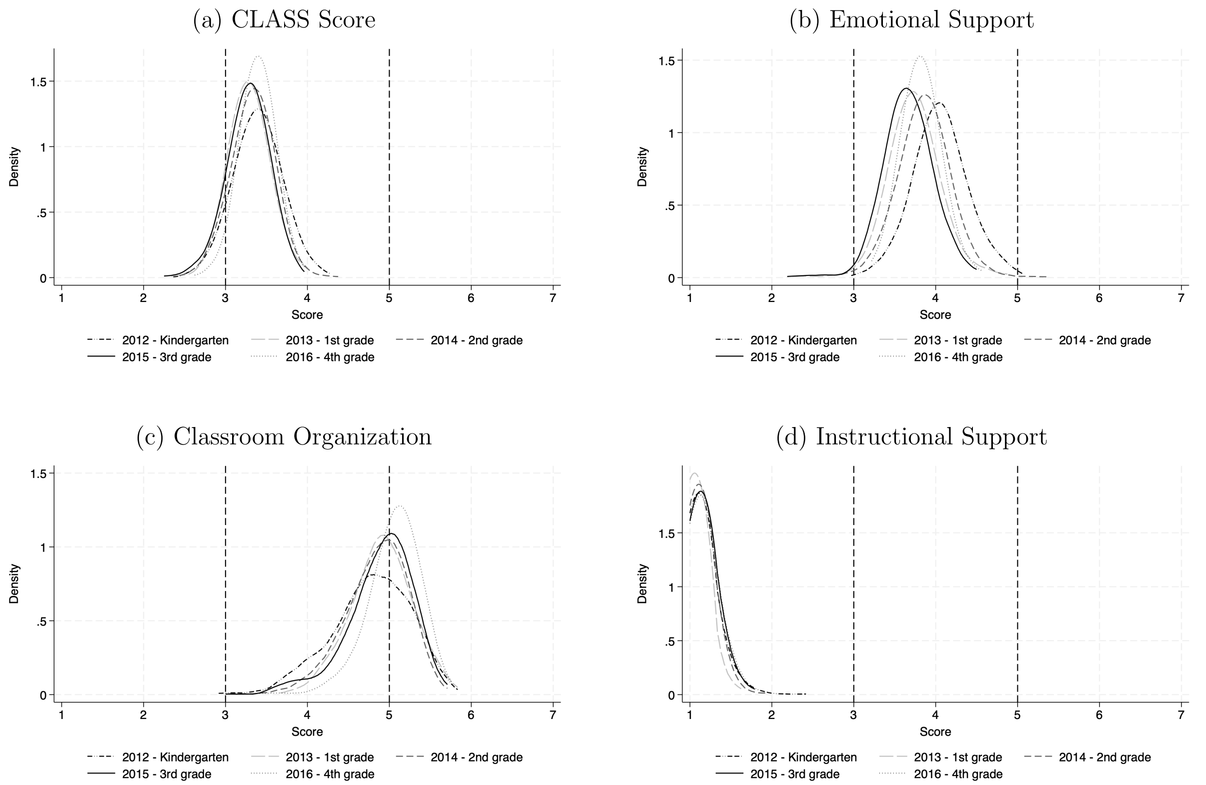
The filming and coding Protocols for the CLASS in Ecuador provide further details on how the process was implemented and how the segments were scored. The authors provide these upon request.

Figure D.1: ClASS Domains and Dimensions



Notes: The figure shows the three domains of the Classroom Assessment Scoring System (CLASS) and the dimensions included in each domain.

Figure D.2: Distribution of CLASS Score and domains



Notes: The figure above presents univariate densities of the distribution of Classroom Assessment Scoring System (CLASS) score and its domains, by grade.

Table D.1: Correlation across CLASS scores

	Total Score	Emotional Support	Classroom Organization	Instructional Support
Total Score	1.000			
Emotional Support	0.881	1.000		
Classroom Organization	0.862	0.563	1.000	
Instructional Support	0.587	0.391	0.398	1.000

Notes: The table presents the results from pairwise correlations between CLASS score and its components collected from Kindergarten to 4th grade. All the correlations are significant at the 1 percent level.

E Extra Robustness Checks

Peer's Parents Education

In this appendix, we estimate the main results in Table 3 using the peer's parents education which has been used in the US context to estimate the effect of high flyers (see Cools et al. (2022)). In particular, these studies use the proportion of leave-one-out proportion of peers with at least one post-college parent. In our context, we are not able to separate college and post-college. For that reason, in Table E.1, we present the results using the leave-one-out proportion of peers with at least one post-secondary parent. The table shows that the coefficients are negative but not significant.

Nevertheless, whether using the proportion of peers with at least one post-secondary parent is appropriate in the context of Ecuador is unclear, given that the proportion of children who have at least one post-secondary parent is smaller. Table E.2 shows that the proportion of children that have at least one post-secondary school is 13.01%. While in the US around 25% of the children have at least one post-college parent (Cools et al., 2022). Given this, in order to define a similar proportion of children as high flyers in both countries, it is necessary to adjust the parents education conditions in Ecuador. Table E.2 shows that around 22% of the children in Ecuador have parents who have finished secondary school.

Table E.3 shows that the results using the proportion of children that both parents finished secondary school are mostly consistent with the results in Table 3 for Executive function. In particular, Executive function skill test scores reduce by 0.014 SD when the proportion increases by one standard deviation. Although math scores do not exhibit significant decreases, this is potentially due to the loss of identifying variation from the smaller sample size. However, it is important to note that the coefficient sign remains the same on all of the estimates.

Table E.1: Effects of Peers' Parents Post-Secondary Education on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of Peers with at least one Post-Secondary Parent	-0.044	-0.106
	(0.058)	(0.081)
Mean of Dependent Variable	0.054	0.038
Treatment Effect of 1SD increase	-0.005	-0.011
Observations	72298	54629
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of peers that have at least one parent with post-secondary education. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table E.2: Proportion of children by parents' education

	Percentage of Children
At least one post-secondary parent	13.01%
Both parents finished secondary school	21.55%

Notes: The table reports the proportion of children whose parents have certain education characteristics in Ecuador.

Table E.3: Effects of Peers' Parents Secondary Education on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of Peers with Parents who finished Secondary	-0.048	-0.111*
	(0.052)	(0.059)
Mean of Dependent Variable	0.054	0.038
Treatment Effect of 1SD increase	-0.007	-0.016
Observations	72298	54629
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of peers that have at least one parent with post-secondary education. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Attrition

In this appendix, we estimate if having a higher proportion of high achievers increases the probability that a child attrits from our sample of school in between grades. Appendix Table E.4 shows the impact of the (leave-one-out) proportion of high achievers on the likelihood of leaving the sample between two consecutive grades. It shows that the children are no more likely to attrit when exposed to a higher proportion of high achievers. Therefore, we do not see evidence of selective attrition. Nevertheless, in Appendix Table E.5, we restrict our sample to the balanced panel of children and estimate the main equation. We find that the results are similar.

Table E.4: Effects of High Achievers on Attrittion

	All grades	2nd grade	3rd grade	4th grade	5th grade	6th grade
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged Prop. of High Achiever Peers	-0.004	-0.001	0.002	-0.037	0.019	-0.000
	(0.011)	(0.031)	(0.013)	(0.041)	(0.017)	(0.015)
Mean of Dependent Variable	0.0187	0.0392	0.00745	0.0338	0.00873	0.00745
Observations	68481	11794	13287	14329	14312	14759
Controls	Yes	Yes	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the lagged leave-one-out proportion of high achievers on the likelihood of being an attritor between t and $t+1$. The first column pools all the grades in a single regression and the remaining columns correspond to a different grade. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table E.5: Effects of High Achievers on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of High Achiever Peers	-0.161**	-0.299***
	(0.080)	(0.100)
Mean of Dependent Variable	0.086	0.063
Treatment Effect of 1SD increase	-0.011	-0.021
Observations	51168	38492
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test score in the previous year or baseline. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Standard errors are corrected for heteroskedasticity and are clustered at the school level.