

UC Santa Barbara

UC Santa Barbara Electronic Theses and Dissertations

Title

Essays in Human Capital Formation

Permalink

<https://escholarship.org/uc/item/2p38g7v7>

ISBN

9798186381563

Author

Fuertes-Segura, Nicolas

Publication Date

2026-06-06

Peer reviewed|Thesis/dissertation

University of California
Santa Barbara

Essays in Human Capital Formation

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Economics

by

Nicolás D Fuertes-Segura

Committee in charge:

Professor Heather Royer, Chair
Professor David Silver
Professor Kelly Bedard

June 2026

The Dissertation of Nicolás D Fuertes-Segura is approved.

Professor David Silver

Professor Kelly Bedard

Professor Heather Royer, Committee Chair

May 2026

Essays in Human Capital Formation

Copyright © 2026

by

Nicolás D Fuertes-Segura

To my family and friends. Thank you for believing in and supporting me.

Acknowledgements

I am grateful to my committee members for their invaluable guidance, patience, and support throughout my PhD journey. I thank Heather Royer, my advisor, for her support from day one of the Ph.D. program and for providing inspiration, encouragement, and strength when I needed it most. I will never forget our many conversations, during which she was always thoughtful and encouraging. I hope to follow in her footsteps in the way she approaches research, teaching, and mentoring younger generations. I also thank Kelly Bedard for her calm and practical perspective and for our many conversations about education systems and how to improve them. Finally, to David Silver for his technical feedback. The three of you have taught me how to be a better economist.

I also thank my colleagues at the Inter-American Development Bank. To Yyannú Cruz-Aguayo, my coauthor and long-time mentor, thank you for helping me understand the importance of connecting research and policymaking and for supporting me since my time at the Bank. Thank you for being so generous with your time and knowledge and for encouraging me to explore new ideas. To Daniel Hernandez, my friend, thank you for your support and for being there for every step of the way, and for reading countless drafts of my papers.

I also thank the broader UC Santa Barbara community. My research has benefited immensely from the incredible faculty, graduate students, and staff in the Department of Economics and the Broom Center for Demography. I am especially grateful to Javier Birchenall, Shelly Lundberg, Alisa Tazhitdinova, and Gonzalo Vazquez-Bare for their support and guidance. I also thank my friends Juan Nicolas, Susana, Sofía, Camila, Matías, Jimena, Jose, and Daniela for making the time in Santa Barbara much more enjoyable and for celebrating every milestone along the way.

Finally, I am deeply grateful to my friends and family. I am deeply thankful to my

mother, Sandra, my grandmother, Maria, and my brother, Tomas, for their unconditional love and support, and for teaching me how to overcome challenges and remain positive. Outside of UCSB, I thank Camila Ayala and Fernando Morales for their friendship. I have greatly valued sharing many drafts with them and am deeply thankful for their encouragement and support throughout this journey.

Curriculum Vitæ

Nicolás D Fuertes-Segura

Education

2026 (Expected)	Ph.D. in Economics, University of California, Santa Barbara
2021	M.A. in Economics, University of California, Santa Barbara
2017	M.A. in Economics, Universidad de Los Andes
2015	B.A. in Economics minor in French Language and Culture, Universidad de Los Andes

Research Fields

Labor Economics, Health Economics, Development Economics

Working Papers

“Friendship Formation and Network Effects: Experimental Evidence from Ecuador,” *Job Market Paper* with Yannú Cruz-Aguayo

“The Asymmetric Effects of High Achiever Peers: Experimental Evidence from Ecuador,” with Yannú Cruz-Aguayo and Carolina Echeverri

“The Impacts of Terrorist Attacks on Teenagers’ Socioemotional Skills and Risky Health Behaviors”

“Medium-Term Impacts of Access to Daycare on School Outcomes: Experimental Evidence from Rio de Janeiro,” with Pedro Carneiro, Sofia Castro, Yannú Cruz-Aguayo, Gregory Elacqua and Norbert Schady

Research Positions

2022	Gretler Fellow for Heather Royer, UC Santa Barbara
2017	Research Assistant for Catherine Rodríguez & Adriana Camacho, Universidad de Los Andes
2015-2017	Research Assistant for Raquel Bernal, Renata Pardo & Lina Castaño, Universidad de Los Andes
2014-2015	Research Assistant for Andrés Moya, Universidad de Los Andes

Professional Experience

Summer 2021-2025	Product External Consultant, Inter-American Development Bank
2017-2020	Consultant, Inter-American Development Bank
2016-2017	Data Center Coordinator, Universidad de Los Andes

Teaching Experience

2023-present	<i>Head Teaching Assistant</i> , Department of Economics, University of California, Santa Barbara Economics 2: Principles of Macroeconomics (multiple quarters); Economics 9: Introduction to Economics, ~450-700 students per quarter
2020-2022	<i>Teaching Assistant</i> , Department of Economics, University of California, Santa Barbara Principles of Microeconomics (multiple quarters); Intermediate Microeconomics I (multiple quarters); Intermediate Macroeconomics (multiple quarters); Introduction to Econometrics (2022)
2015-2017	<i>Teaching Assistant</i> , Department of Economics, Universidad de Los Andes The Economics of Risky Health Behaviors (Graduate); Principles of Macroeconomics, Applied ELCA Research Seminar; Measurement and Accounting

Honors and Awards

2025	ASHEcon "Health and Development" Program Chair Award
2025	Doctoral Student Travel Grant, UCSB Academic Senate
2025	Conference Travel Grant, Graduate Student Association, UCSB
2024	APPAM Equity & Inclusion Fellowship
2024	Research Quarter Fellowship, UCSB
2024	Doctoral Student Travel Grant, UCSB Academic Senate
2024	Conference Travel Grant, Graduate Student Association, UCSB
2023	ASHEcon Diversity Scholarship
2022	Gretler Fellowship, UCSB
2021	Distinction in Ph.D. Preliminary Exam, Econometrics, UCSB

Conference Presentations and Seminars

2026	Ridge Forum (Barbados), Harvard University Center for Education Policy Research
2025	APPAM (Seattle), ASHEcon (Nashville), Southern California Graduate Conference in Applied Economics (Los Angeles), Western Economic Association Annual Conference (San Francisco), UCSB Applied Microeconomics Lunch
2024	APPAM (National Harbor, Maryland), ASHEcon (San Diego), UCSB Applied Microeconomics Lunch
2023	Southern California Graduate Conference in Applied Economics (Los Angeles)

Professional Service

	<i>Referee:</i> American Economic Journal: Economic Policy
2024–2026	Abstract Reviewer, ASHEcon Annual Conference
2022–2025	Co-Chair of the Mentorship Program First-Year Ph.D. Students in Economics, University of California, Santa Barbara
2021–2025	Mentor, First-Year Ph.D. Students in Economics, University of California, Santa Barbara
2021–2024	Organizing Committee Member for the Southern California Graduate Conference in Applied Economics

Computer Skills

Stata, L^AT_EX, Matlab, R

Languages

English (native proficiency), Spanish (native), French (intermediate), Portuguese (basic)

Permissions and Attributions

1. The content of chapter 1 is the result of a collaboration with Yyannú Cruz-Aguayo.
2. The content of chapter 2 is the result of a collaboration with Yyannú Cruz-Aguayo and Carolina Echeverri.

Abstract

Essays in Human Capital Formation

by

Nicolás D Fuertes-Segura

This dissertation examines how social environments and early-life shocks shape cognitive, socioemotional, and behavioral outcomes in developing countries. Across three chapters, I combine experimental data, panel data, survey data, and quasi-experimental methods.

The first chapter, coauthored with Yyannú Cruz-Aguayo, studies the causal impact of friendships on children's outcomes in elementary schools in Ecuador. While the literature on peer effects typically focuses on broad measures of peer composition, such as average ability or class characteristics, friendships represent a more direct and behaviorally relevant form of interaction. Furthermore, little is known about whether friendships help or hurt learning and well-being. Friendships can provide emotional support and foster motivation, but they may also distract from academic activities or reinforce negative behaviors. We use a unique randomized experiment in which students are randomly assigned to classrooms in every grade from Kindergarten to 6th grade, combined with detailed network data. We show that being assigned to a classroom with more friends has contrasting effects across domains: it reduces cognitive outcomes, such as math and executive function, while improving emotional well-being. These findings highlight that friendships hurt academic performance but improve socioemotional outcomes. We further show that these negative effects are driven by the number of disruptive friends and are concentrated among male students. Overall, these findings suggest that friendship networks may affect learning through a distraction or lack of attention channel.

The second chapter, coauthored with Yyannú Cruz-Aguayo and Carolina Echeverri, examines how exposure to high achievers in the classroom affects cognitive and non-cognitive skills. Using the same experimental setting, we exploit random variation in the proportion of high achievers across classrooms. We find that exposure to high achievers reduces math and executive function scores, with stronger effects among top-performing students. The results suggest that heightened competition, greater peer interaction, and increased salience of relative class rank play an important role in shaping outcomes. In addition, we show that higher-quality teachers can mitigate these negative effects, highlighting the interaction between peer composition and classroom management.

The third chapter studies the impact of exposure to violence during childhood (0-9 years old) on socioemotional skills and risky health behaviors in adolescence. Using data from the Colombian Longitudinal Survey and municipality-level administrative data, and exploiting within-municipality variation across cohorts in exposure to terrorist attacks during childhood, I estimate the effects of violence on adolescent outcomes. The results show that exposure to conflict affects socioemotional development and increases engagement in risky behaviors, particularly drinking alcohol. These findings contribute to the literature on the potential long-term consequences of violence and highlight the importance of early-life conditions in shaping behavioral and health outcomes during adolescence.

JEL Codes: I21, J16, I12, J13, D85, Z13, D10, F51

Contents

Curriculum Vitae	vii
Permissions and Attributions	x
Abstract	xi
1 Friendship Formation and Network Effects: Experimental Evidence from Ecuador	1
1.1 Introduction	1
1.2 Setting and Data	7
1.3 Empirical Strategy	22
1.4 Results	30
1.5 Conclusion	53
2 The Asymmetric Effects of High Achiever Peers: Experimental Evidence from Ecuador	56
2.1 Introduction	56
2.2 Setting and Data	61
2.3 Empirical Strategy	75
2.4 Results	81
2.5 Conclusion	99
3 The Impacts of Terrorist Attacks on Socioemotional Skills and Risky Health Behaviors	101
3.1 Introduction	101
3.2 Background and Prior Literature	107
3.3 Data	112
3.4 Empirical Strategy	120
3.5 Results	128
3.6 Conclusion	148

A	Appendix for Chapter 1	150
A.1	Supplemental Figures and Tables	151
A.2	Homophily Descriptives	159
A.3	Details on the Tests and IRT	161
A.4	Application of the CLASS in Ecuador	163
B	Appendix for Chapter 2	166
B.1	Supplemental Figures and Tables	167
B.2	Randomization details	178
B.3	Details on the Tests and IRT	179
B.4	Application of the CLASS in Ecuador	182
B.5	Extra Robustness Checks	186
C	Appendix for Chapter 3	191
C.1	Supplemental Figures and Tables	192
C.2	Details on Sample Selection	203
C.3	Extra estimates on Socioemotional Skills Components	204

Chapter 1

Friendship Formation and Network Effects: Experimental Evidence from Ecuador

1.1 Introduction

Social networks play a central role in shaping behavior, performance, and decision-making across education, workplaces, and broader social contexts, with both short- and long-term consequences for individual and collective outcomes (Granovetter, 1973).¹ In education, previous studies have analyzed the role of social networks primarily in college (Barrios-Fernández, 2022; Zhu, 2025; Brade, 2025) and secondary school settings (Bramoullé et al., 2009; Lavy and Sand, 2019; Fletcher et al., 2020). However, given the long-term consequences of early childhood development, it is critical to understand how friendship networks formed early in life shape students' outcomes. Identifying the

¹See Sacerdote (2001); Carrell et al. (2013); Lavy and Sand (2019); Barrios-Fernández (2022) for educational outcomes; Bayer et al. (2008); Hellerstein et al. (2011); Hensvik and Skans (2016); Park (2019); Zhu (2025) for labor market outcomes; and Chetty et al. (2022) for social and economic mobility.

mechanisms underlying these effects can inform policies that enhance children’s outcomes, particularly in developing countries where resource constraints and institutional incentives differ substantially.²

The impact of social networks is both theoretically ambiguous and empirically mixed. Although some studies find positive effects due to increased motivation and social pressure (Babcock et al., 2015) or greater cooperation and social satisfaction (Lavy and Sand, 2019), others find negative effects if they are used for socializing or in groups with less able friends (Bandiera et al., 2010). Moreover, these effects can be positive or negative depending on the type of network analyzed (e.g., reciprocal friendship or rejecter) (Lavy and Sand, 2019). In this paper, we study how friendship networks affect both cognitive and non-cognitive skills during elementary school, an early and critical period when peer exposure is intense and in which cognitive and non-cognitive skills are strong predictors of long-term outcomes (Moffitt et al., 2011; Heckman et al., 2006, 2013).

In this paper, we use data from a unique randomized experiment involving the 2012 entering Kindergarten cohort across 202 elementary schools in Ecuador. In these schools, at the beginning of each grade, from Kindergarten through sixth grade, students and teachers were randomly assigned to classrooms within their schools, each of which had at least two classrooms per grade.³ Compliance with the random assignment was nearly perfect (98.9% on average). Thus, students who remained in the same school experienced seven exogenous, orthogonal peer groups.⁴ This unique random assignment over several years lets us analyze how classroom composition affects friendship formation

²For example, over the past 25 years, developing countries have invested heavily in education. In Latin American and the Caribbean the average government expenditure in education has been around 4.5% of the GDP since 2000 and during that same period primary completion rates went from 90 to 96 percent. All data were obtained from the World Development Indicators (<https://datatopics.worldbank.org/world-development-indicators/>) on August 15th, 2025.

³Each grade is taught by a completely different set of teachers, as is customary in Ecuador

⁴Classroom and teacher quality vary but are orthogonal to peer and friend quality, as teachers were also randomly assigned within schools and grades.

and persistence, and the impact of these friendship networks on cognitive (math and executive function) and non-cognitive (depression, self-esteem, growth mindset, and grit) skills during elementary school. A major challenge for studying friendship formation, persistence, and their effects is the endogeneity of friendship choice. This design addresses the issue directly, as students are randomly assigned to classrooms each grade.

For this cohort, we observe rich longitudinal data spanning seven consecutive years, including baseline characteristics, test scores, non-cognitive skills, teacher characteristics, and teacher quality indicators. Furthermore, we elicited students' friendship networks at the end of both third and sixth grades. In third grade, each student could nominate up to three friends from the same school, and in sixth grade, up to five. We use these data to construct directed links (i.e., student i nominates student j , regardless of whether student j reciprocates), reciprocal friendships (i.e., both student i and student j nominate each other), followers (i.e., those who nominated a student but were not reciprocally nominated) and rejecters (i.e., those who are nominated by a classmate but do not reciprocate) in both grades within each school. Based on the third-grade network, we construct, for each student, the number of directed links, reciprocal friends, followers, and rejecters at the school and classroom levels for grades 3–6, using the classroom composition of each grade.

First, we examine how the timing and duration of classroom exposure shape friendship formation in third grade. We find that, after controlling for school fixed effects, pairs of students randomly assigned to the same classroom are significantly more likely to nominate each other as friends than pairs who were never classmates. Conditional on being randomly assigned in third grade, longer cumulative exposure further increases the likelihood that a friendship forms, suggesting that repeated interaction fosters social ties. We also provide descriptive evidence on the extent of homophily: children are more likely to report being friends with peers of similar ability and the same gender, whereas similarity in maternal education plays a more limited role. Additionally, disruptive students often

form friendships with other disruptive students.

Second, we estimate how the number of classroom friends affects cognitive (math and executive function) and non-cognitive skills (depression, self-esteem, growth mindset, and grit). In particular, we estimate the effect of the students' friendship network in the classroom after the random assignment in each grade from fourth to sixth grade. Our main specification includes school-by-grade fixed effects and the number of friends in the school. This allows us to compare children who attended the same school and grade, and have the same number of school friends, but were randomly assigned to different classrooms. Therefore, conditional on the school and on the number of friends a student has at the school, the number of friends a student has in their classroom from fourth to sixth grade should be random. We find that having one additional directed friend in the classroom reduces math test scores by 0.006 SD, and executive function, a set of basic self-regulatory skills, by 0.023 SD. These results are robust across various specifications (detailed in section 1.4.2). Alongside these reductions in cognitive skills, we find that having one additional directed friend in the classroom reduces depression in sixth grade. This pattern suggests that the effects are not driven by lower happiness or self-confidence. Overall, the findings highlight the importance of accounting for a student's friendship networks when evaluating how classroom environments affect cognitive development and how educational policies affect outcomes.

There are notable gender differences. Indeed, the negative effects are concentrated among male students and we find no effects on female students. These results highlight the mixed evidence on whether friendship networks improve performance. In particular, given the differences by gender, it suggests that male students are potentially negatively affected by the use of the network for socializing or if in groups with less able friends (Bandiera et al., 2010). Furthermore, consistent with previous findings (Lavy et al., 2012), we find that the number of rejecters negatively affects cognitive skills, which suggests that

a behavioral response make take place by lowering happiness and reducing effort.

While we cannot conclusively distinguish between different mechanisms, our findings suggest that the negative effects of the number of friends in the classroom likely operate through a distraction or lack of attention channel. First, when examining friend characteristics, we find that the negative effects for male students are driven by the number of directed disruptive friends in the classroom. Second, the negative effects of disruptive friends are concentrated among high-ability students. Therefore, these results suggest that when high-ability male students are randomly assigned to classrooms with less-able friends or when male students are assigned to the same classroom as disruptive friends, classroom behavior deteriorates, thereby negatively affecting learning.

A potential concern in estimating network effects under random classroom assignment is the weak variation in group composition as classroom size increases (Angrist, 2014). In our setting, identification relies on residual variation in the number of friends in the classroom due to finite-sample imbalances in classroom assignments. To assess this variation, we examine whether the number of classroom friends a student has follows a hypergeometric distribution with school and student-specific parameters. Only 2.01% of classrooms have a number of friends in the classroom that lies outside the 95% confidence interval implied by this distribution. Furthermore, on average, the observed mean and standard deviation match the hypergeometric theoretical expectation. Additional diagnostic analyses, described in Section 1.3.2, further support the validity of our empirical strategy.

This paper contributes to the literature in four ways. First, we extend the literature on the impacts of friendship networks beyond labor market outcomes (Bandiera et al., 2010), crime (Billings et al., 2014), health outcomes (Fletcher and Ross, 2018), and achievement (Fletcher et al., 2020; Lavy and Sand, 2019; Brade, 2025), by studying effects on executive function and non-cognitive skills during elementary school, a period in which

children’s environments can have long-term effects (Heckman et al., 2013). We leverage a unique experiment with repeated random assignment and nearly perfect compliance to estimate the causal effects of the number of friends in the classroom and to document their non-linear impacts. Moreover, our rich data on teacher quality allow us to analyze how friendship networks interact with other classroom inputs. A novel aspect relative to prior work is that we analyze how friends’ characteristics mediate how the network affects learning and likely operate through a distraction or lack of attention channel.

Second, it adds to the evidence on network formation. While previous studies have analyzed network formation with structural models (Currarini et al., 2009; Goldsmith-Pinkham and Imbens, 2013; Graham, 2017; Johnsson and Moon, 2021; Jackson et al., 2025), this study leverages the random assignment of children to classrooms to estimate the causal effects of classroom exposure on friendship formation. Furthermore, a novel aspect of our study is that we leverage a unique longitudinal experiment that elicits friendship networks in two different grades, which enables us to study friendship persistence.

Third, we contribute to the literature on homophily in friendships (Verbrugge, 1977; McPherson et al., 2001; Currarini et al., 2009; Carrell et al., 2013; Chetty et al., 2022). Previous studies have examined homophily by race, ethnicity, and gender (Shrum et al., 1988; Jackson et al., 2025), malleable characteristics (e.g., altruism) (Jackson et al., 2025), and substance use (Pearson et al., 2006), primarily among adolescents and college students. We provide new descriptive evidence on elementary school children, for whom homophily may be less ingrained. Furthermore, using our rich data, we extend this literature by analyzing homophily based on a new set of characteristics: disruptiveness and prior PPVT scores.

Finally, our findings inform student-grouping policies, especially in developing countries that experience larger class sizes, tighter resource constraints, and different institutional

settings.⁵ We provide new evidence from elementary schools in Ecuador, underscoring the importance of accounting for context-specific network dynamics and institutional settings when designing grouping policies to improve student outcomes.

The remainder of the paper is organized as follows. Information about Ecuador, the experimental setting, and data can be found in Section 1.2. Section 1.3 outlines the empirical strategy. Section 1.4 provides the main results, the dynamics of the effects, and heterogeneity analysis. Conclusions and policy discussion can be found in Section 1.5.

1.2 Setting and Data

1.2.1 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 ages 5 to 14. The education system is divided into elementary school (Kindergarten through 6th grade), middle school (7th through 9th grades), and high school (10th through 12th grades). The school year follows a dual-calendar system. In the coastal region, it runs from May to February (similar to most countries in the Southern Hemisphere and many in Latin America), while in the highlands and Oriente (eastern) regions, it runs from September to June (similar to most countries in the Northern Hemisphere).⁶

Approximately 4.4 million children were enrolled in the education system during the 2012–2013 school year, with 78% attending public schools.⁷ Ecuador has made significant progress in expanding access to education. In 2013, the primary school net enrollment

⁵See Sacerdote (2011, 2014) for a review of previous studies.

⁶According to the 2010 census, 53% of the population lived on the coast, 42% lived in the highlands, and 5% in the Oriente region.

⁷Data obtained from <https://educacion.gob.ec/datos-abiertos/> on November 4th, 2024.

rate was 95%, with a completion rate of 97%.⁸ 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 primary school children achieved the minimum proficiency level in mathematics, with a slight gender gap—49.8% for males compared to 46.7% for females.⁹ 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 both the quality of the education system and the instruction provided to students.¹⁰ However, key challenges persist. For example, the pupil-qualified teacher ratio remained stable at around 27.5 during this period.¹¹

1.2.2 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.¹² Each school had at least two classrooms per

⁸The net enrollment rate is the fraction of children of school age who are enrolled in school, while the completion rate corresponds to the percentage of children who have finished the last grade of primary.

⁹This corresponds to the percentage of children in primary who exceed the minimum proficiency level (MPL) which is the pre-defined proficiency level of basic knowledge in mathematics measured through learning assessments. All data were obtained from the World Development Indicators (<https://datatopics.worldbank.org/world-development-indicators/>) on November 4th, 2024.

¹⁰This corresponds to the fraction of teachers with the minimum academic qualifications required for teaching in primary school.

¹¹All data were obtained from the World Development Indicators (<https://datatopics.worldbank.org/world-development-indicators/>)

¹²These schools are a random sample of all public schools with at least two Kindergarten classrooms in Ecuador’s coastal region in 2012. See Araujo et al. (2016) for more details on the school selection. Moreover, they show that the characteristics of students and teachers in the sample closely resemble those in a nationally representative sample of schools in Ecuador.

grade, with most having exactly two. In the 2012 school year, an incoming cohort of children was randomly assigned to Kindergarten classrooms within each school. These children were then randomly reassigned to classrooms in subsequent grades through 6th grade in 2018. Compliance with the assignment rules was very high, averaging 98.9% (Carneiro et al., 2025b). As a result, children who remained in the same school throughout elementary school were exposed to seven exogenous, orthogonal shocks to classroom composition.¹³

Previous studies have used this experiment to estimate teacher value-added and classroom effects (Araujo et al., 2016) and to study gender gaps and heterogeneous responses to the quality of classrooms, teachers, and peers (Carneiro et al., 2020). These data have also been used to analyze the impact of various classroom inputs on test scores, including relative classroom rank (Carneiro et al., 2025b), peer composition (Carneiro et al., 2025a; Fuertes-Segura et al., 2025), and test-screened tenured teachers (Araujo et al., 2025). Our study departs from the previous studies by combining the random assignment with the elicitation of friendship networks in two grades to identify how exposure shapes friendship formation and persistence, and how classroom friends affect cognitive and non-cognitive outcomes.

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).¹⁴ Using this dataset, Carneiro et al. (2025b) test the success of the random assignment by applying the method proposed by Jochmans (2023).

¹³It is worth noting that random reassignment may separate students from their friends. However, out-of-class friendship ties are orthogonal to peer composition given the random assignment.

¹⁴We 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, children were ordered by their last and first names and 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.

This method 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 . They show that they cannot reject the null hypothesis that there is no correlation between child i 's achievement and that of her classroom peers, confirming that the random assignment was successful. Together, these features provide repeated, exogenous variation in classroom peers and rich measurement, which we leverage to study friendship formation, persistence, and their effects on later skills.

1.2.3 Data

Child data

At the beginning of Kindergarten, we collected baseline data on maternal education, household wealth, preschool attendance, and children's vocabulary skills.¹⁵ 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 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).¹⁶ We aggregated correct math responses for each component separately using Item Response Theory (IRT) and then computed the total math score, with each component receiving equal weight.

Additionally, we collected data on executive function for each grade from Kindergarten

¹⁵To 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.

¹⁶The 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 0, and the right end 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)).

to 4th grade. Executive function encompasses a set of self-regulatory skills involving various regions of the brain, particularly the prefrontal cortex. It is generally divided into three domains: working memory and attention, inhibitory control, and cognitive flexibility.¹⁷ 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 solve math problems, among other classroom tasks. More importantly, these skills have been shown to predict long-term outcomes in adulthood related to labor market success, health and crime (Moffitt et al., 2011). We compute scores for each of the three domains, as well as an overall score.¹⁸

In 6th grade, we collected data on child depression, self-esteem, growth mindset, and grit. To measure depression, we used the Patient-Reported Outcomes Measurement Information System (PROMIS) Depression Scale for children aged 11-17, 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 chose 10 questions from the Dweck "Mindset Quiz"; growth mindset refers to the belief that intelligence is malleable and can be developed with effort (Blackwell et al., 2007). Finally, to measure grit, we adapted four questions from the 8-item Grit Scale for children (Duckworth and Quinn, 2009); grit refers to the ability

¹⁷Working 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 refers to 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.

¹⁸Unlike 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.

to persevere at a given task. For each outcome, we aggregated responses using factor analysis. Further details on the child assessments and the IRT are provided in Appendix A.3.

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

Finally, at the end of each grade, we asked teachers to identify five students with the most severe behavioral problems. Therefore, our homophily analysis focuses on the behaviors that teachers believe disrupt learning and captures students with persistent behavioral problems. To mitigate concerns about subjectivity in these rankings, we classify a student as disruptive in third grade if at least 50% of the teachers who observed the student in previous grades identified the student as such.¹⁹ By aggregating multiple, independent evaluations, this procedure reduces measurement error from any single teacher assessment or from individual teacher behavior toward particular students, and provides a stronger signal of student behavior.

Friends Information

At the end of both third and sixth grades, we collected data on friendship networks. In third grade, each student could nominate up to three friends from the same school, and

¹⁹For example, a child identified as having severe behavioral problems by three former teachers would be considered disruptive. Importantly, teachers appear only once in our sample and do not teach in future grades.

in sixth grade, up to five. We use these data to construct directed links and reciprocal friendships in both grades within each school, with on average 87 students. We define a directed link as one where student i nominates student j , regardless of whether student j reciprocates. For a school with N_s students, the number of potential directed links is $N_s(N_s - 1)$. A reciprocal friendship occurs when both student i and student j nominate each other. The number of potential reciprocal friendships is $N_s(N_s - 1)/2$, since we do not include both ij and ji . Directed links are further classified into two groups: (1) followers, those who nominated a student but were not reciprocally nominated, and (2) rejecters, those who are nominated by a classmate but did not reciprocate.

We rely on both directed and reciprocal links because each provides distinct advantages. Directed links may capture cases in which students nominate peers they aspire to befriend, reflecting an aspirational preference channel that increases the overall level of friendship ties in the network. Furthermore, given that students are only allowed to nominate up to 3 or 5 students, relying only on reciprocal links would miss some genuine friendships. For instance, if a student had a fourth or fifth friend in third grade, these connections would not appear in our data. By contrast, reciprocal friendships are more likely to represent stronger, more established ties (Goodreau et al., 2009). Moreover, these links capture "best friends" rather than acquaintances (Fletcher et al., 2020).

Finally, to estimate the effects of friendship links on test scores, we calculate for each student the number of directed links, followers, rejecters, and reciprocal friends both at the school and classroom levels in third grade. Since classroom composition changes in fourth through sixth grades due to the random assignment process, we also calculate these four types of links for each student at both the school and classroom levels for grades 4–6.

Teacher and Classroom data

We use the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2015) to measure teacher behaviors based on recordings of classroom interactions over the course of an entire school day. The CLASS measures teacher behavior across three domains: Emotional Support, Classroom Organization, and Instructional Support.²⁰ Within each domain, there are three to four CLASS dimensions, each rated on a scale of 1 to 7. The CLASS protocol provides coders with specific guidance for assigning scores, categorizing them as "low" (1–2), "medium" (3–5), or "high" (6–7). The behaviors coders assess within each dimension are clearly defined and highly 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 informed of the filming only on the day itself, ensuring that the videos captured typical, planned interactions with students. Each classroom has only one teacher that teaches the same class during the calendar year, without a classroom aide, who was responsible for all academic subjects (excluding physical education, 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.²¹ Further

²⁰Emotional Support measures how teachers promote a positive classroom climate and respond to students' emotional needs. Classroom Organization refers to the routines and procedures teachers use to manage students' behavior, time, and attention. Instructional Support evaluates how effectively teachers promote students' thinking through their implementation of the curriculum.

²¹In cases where the two coders provided significantly different scores, a third coder was assigned to evaluate the teacher. The inter-coder reliability was approximately 0.92, suggesting that measurement

details on the CLASS dimensions, the scoring protocol, and its application in Ecuador are provided in Appendix A.4.

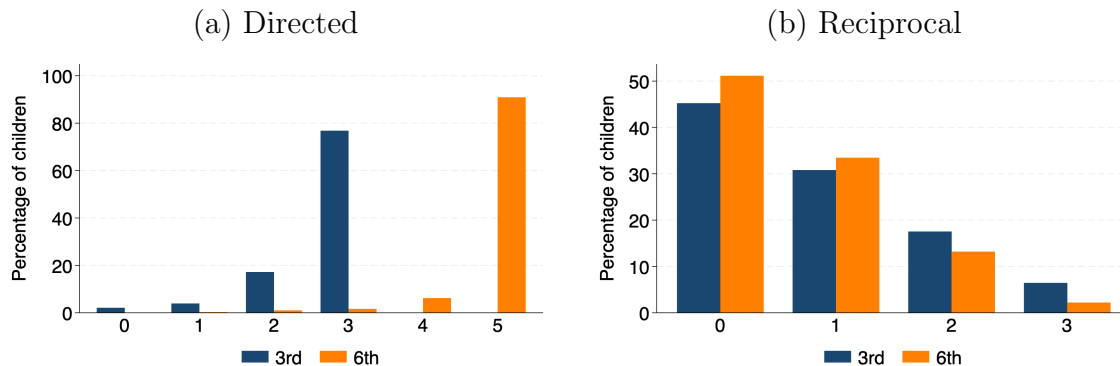
One disadvantage of the CLASS is that it is not a cardinal scale, which means it cannot be used directly as an input in the production function, unlike value-added, experience, and other teacher characteristics. However, previous studies have found that teacher behaviors, measured by the CLASS, are associated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

1.2.4 Descriptive Statistics and Preliminary Checks

Figure 1.1 shows the distribution of directed links and reciprocal friends that a student has in third and sixth grades, separately by grade. The Figure shows that around 80% of students in third grade and 90% in sixth grade report the maximum number of friends possible. This pattern suggests that most students probably experienced a ceiling effect, as they named all possible friends, and we are unable to observe whether they would have named additional friends if given the opportunity to list more. Regarding reciprocal friendships, the Figure shows that approximately half of the students do not have any reciprocal friends. Notably, despite being able to nominate more students in sixth grade, the distribution of reciprocal friends remains concentrated at zero or one, and in our data no student has more than three reciprocal friends. This pattern indicates that reciprocal links identify a narrower set of stronger ties. Finally, the ceiling effect is likely smaller for the measurement of reciprocal friendships.

error due to from substantial differences across coders was relatively low.

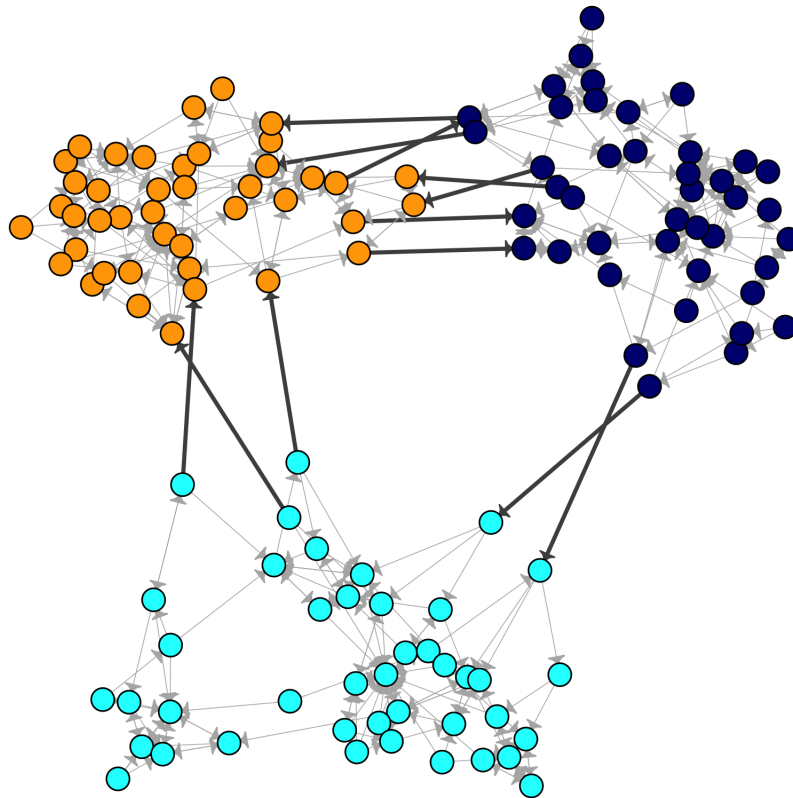
Figure 1.1: Distribution of Friends per Grade



Notes: The figure above represents the distribution of the number of directed friendship links and reciprocal friends in the 3rd and 6th grades, respectively. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other.

Figure 1.2 shows the friendship network for one school in third grade. The Figure presents descriptive patterns related to how friendships form. Each of the three colors represents one of the three classrooms within the school, and each dot represents a student. The light-gray lines represent links between students in the same classroom, and the dark-gray lines represent links between students in different classrooms. First, the Figure shows the importance of the third-grade classroom, as most links occur within classrooms, and very few links occur across classrooms. Indeed, less than 4% of the links in this school are among students in different classrooms. Second, the arrows point to the student who was nominated. The figure shows that reciprocal friendships, those with arrows at both ends, are less common (only 35% of the links in this school are reciprocal).

Figure 1.2: Friendship Network in 3rd grade



Notes: The figure above represents the directed and reciprocal links for a given school in 3rd grade. The three colors represent the three classrooms within the school. Directed links have a single arrow, while reciprocal links have arrows at both ends. Darker arrows indicate links between children in different classrooms. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other.

Table 1.1, Panel A provides summary statistics for the children in the sample. On average, children were eight years old on the first day of third grade, with 62% having attended preschool, and half being girls. The average score on the baseline vocabulary skills (PPVT) test was 83 points, ranging from 55 to 145. Thirteen percent of children were classified as disruptive by their previous teachers. Panel B presents information on

household and parental characteristics. The average age of mothers was 30 years and 34 years for fathers. Regarding maternal education, 39% of the children had mothers with some primary education, 31% with some secondary education and 9% with post-secondary education. Finally, regarding household conditions, 85% reported having piped water, and 46% had a toilet at home.

Table 1.1, Panel C presents information about the friendship network in third grade. On average, students have 2.69 directed links, while reciprocal friends are less frequent, averaging 0.85. Notably, the average number of friends at the classroom level is slightly lower than at the school level, which is consistent with the descriptive pattern discussed in the previous figure. Panel D shows the number of directed links and reciprocal friends for grades fourth to sixth based on the friendship network elicited in third grade. For these counts, we check which third-grade friends remain in the same school and, among them, which share the student's classroom in that grade. The panel shows a decline in both directed links and reciprocal friends at the school and classroom levels, with a larger decrease in classroom friendships. This is expected, as students may be assigned to a different classroom than their friends due to the random classroom assignment process. We rely on this variation in the number of friendships in the classroom, which is random conditional on the number of friends in the school and the school fixed effects.

Table 1.1: Child, Parents, and Friends Network Characteristics

	Mean	SD	Median	Min	Max
A. Child characteristics					
Age of child in 2015	8.63	0.36	8.61	7.21	13.21
Sex (1 = Female)	0.50	0.50	1.00	0.00	1.00
Receptive vocabulary score (PPVT)	83.30	15.76	82.00	55.00	145.00
Lagged Math Test Score	0.00	1.00	0.02	-3.83	3.58
Depression	-0.00	1.00	0.08	-4.17	1.55
Self-esteem	-0.00	1.00	-0.02	-4.76	1.91
Growth Mindset	-0.00	1.00	0.09	-4.79	2.19
Grit	0.00	1.00	0.02	-4.43	1.79
Disruptive	0.11	0.32	0.00	0.00	1.00
Proportion who attended preschool	0.62	0.49	1.00	0.00	1.00
B. Household characteristics					
Maternal Education					
Some Primary	0.39	0.01	0.40	0.38	1.00
Some Secondary	0.51	0.01	0.51	0.00	0.52
Post Secondary	0.09	0.00	0.09	0.00	0.10
Mother's age	30.40	6.58	30.00	5.00	87.00
Father's age	34.73	7.94	33.00	5.00	99.00
Household has piped water in home	0.85	0.36	1.00	0.00	1.00
Household has flush toilet in home	0.46	0.50	0.00	0.00	1.00
C. Friendship Network in 3rd grade					
No. of Directed Friends in School	2.69	0.65	3.00	0.00	3.00
No. of Directed Friends in Classroom	2.58	0.74	3.00	0.00	3.00
No. of Reciprocal Friends in School	0.85	0.93	1.00	0.00	3.00
No. of Reciprocal Friends in Classroom	0.84	0.93	1.00	0.00	3.00
D. Friendship Network in 4th-6th grade					
No. of Directed Friends in School	2.35	0.86	3.00	0.00	3.00
No. of Directed Friends in Classroom	0.98	0.87	1.00	0.00	3.00
No. of Reciprocal Friends in School	0.75	0.88	1.00	0.00	3.00
No. of Reciprocal Friends in Classroom	0.32	0.57	0.00	0.00	3.00

Notes: The Table reports summary statistics of the children in the sample. It includes children's characteristics, household's characteristics, and those of their friendship networks in the school.

We analyze homophily in friendship links along two separate dimensions: gender and maternal education. First, we examine how variation in the number of directed links across groups can affect the observed number of homophilous links.²² Table 1.2 reports the Coleman homophily index, which compares the observed number of within-group directed links to the expected number under random assignment (Coleman, 1958). In this index, a value of 1 represents perfect homophily, while 0 indicates no sorting beyond random assignment. For gender homophily, the index is high in third grade and declines by sixth grade as students start interacting more with opposite-gender classmates, consistent with the literature that early elementary school friendships are largely same-gender (McPherson et al., 2001; Garrote et al., 2023). For maternal education homophily, the index is small but positive, indicating weaker sorting on this trait. Furthermore, we explore homophily by maternal education separately for female and male students. We find that female students exhibit more education-based homophily when their mothers have post-secondary education, while male students exhibit less homophily when their mothers have post-secondary education.

²²For example, even if friendships are formed randomly, the expected fraction of homophilous friendships is not $\sum_{i=1}^n p_i^2$ when different groups have different numbers of friends. Instead, it becomes $\frac{\sum_{i=1}^n p_i^2 k_i}{\sum_{i=1}^n p_i k_i}$ where p_i is the fraction of students with characteristic i in the sample and k_i is the number of nominations made by students with characteristic i

Table 1.2: Coleman Homophily Index for Directed Links

	3rd grade			6th grade		
	Share same group	Group share	Index	Share same group	Group share	Index
A. Child's Gender						
Female	90.40	48.30	0.814	81.58	48.53	0.642
Male	87.75	51.70	0.746	81.05	51.47	0.610
B. Maternal Education						
<i>Female</i>						
Some Primary	45.55	40.16	0.090	46.95	40.48	0.109
Some Secondary	55.32	50.95	0.089	55.70	51.06	0.095
Post Secondary	19.97	8.89	0.122	14.56	8.45	0.067
<i>Male</i>						
Some Primary	47.19	38.44	0.142	44.44	38.80	0.092
Some Secondary	57.60	52.03	0.116	58.69	52.45	0.131
Post Secondary	14.87	9.53	0.059	15.58	8.75	0.075

Notes: The Table reports statistics about homophily in directed links. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. The "Share same group" column presents the mean of the percentage of directed friends from the same gender or same maternal education as their own. The "Group share" column provides the distribution of students by gender and maternal education in the full sample. The "Index" shows the Coleman Homophily Index, which compares the observed number of directed links within groups to the expected number of links under random assignment (Coleman, 1958). In this index, a value of 1 represents perfect homophily, while 0 indicates no sorting due to random assignment.

Furthermore, in Appendix A.2 we provide additional descriptive analysis of homophily. First, Appendix Table A.3 shows that there is substantial within-gender linking; by contrast, the analogous pattern for maternal education is present but weaker. It also shows that these patterns reflect homophilous preferences rather than mechanical composition effects. Second, Appendix Table A.4 compares the expected and observed percentages of homophilous friendships. It shows that the observed share exceeds the expected share. For instance, for gender, we would expect approximately 50% of links to be between students of the same gender, but we observe around 97%. These results are consistent with earlier descriptive findings and suggest that homophily is more pronounced by gender than by maternal education.

1.3 Empirical Strategy

1.3.1 Friendship Formation and Persistence Model

We study how the timing (i.e., which grade) and length (i.e., how many grades) of within-school classroom exposure affect friendship formation in third grade and the persistence of those links to sixth grade. In particular, we leverage variation from the random assignment of students to classrooms. For this reason, we construct binary indicators of whether there is a directed link and whether there is a reciprocal friendship between students i and j in both grades, as well as binary indicators for whether (i, j) were in the same classroom in each grade from Kindergarten to third grade for the formation analysis and from fourth grade to sixth grade for the persistence analysis. With these data, we can investigate the effects of being randomly assigned to the same classroom and the length of exposure on friendship formation and persistence. Formally, we estimate

the following dyadic link formation model:

$$y_{ijs} = \alpha + \sum_{t=0}^3 \beta_t SC_{ijst} + \delta_s + \mu_{ijs} \quad (1)$$

where y_{ijs} is a binary indicator that denotes whether there is a directed link or reciprocal friendship between student i and student j at school s in third or sixth grade. SC_{ijst} is a binary indicator that equals one if students i and j were randomly assigned to the same classroom at school s in grade t , where $t = 0$ is Kindergarten. δ_s is a school fixed effect. We cluster standard errors at the school level, which allows the probability of students becoming friends or the friendship persisting in different classrooms to be correlated within schools. The parameters of interest, β_t , capture the impact of being randomly assigned to the same classroom in grade t on the probability that students i and j form a directed or reciprocal link in third grade and on the probability that the link persists to sixth grade.

Previous literature has shown that homophily, the tendency of individuals to associate with others who share similar demographics and background characteristics, plays a central role in friendship formation and persistence (McPherson et al., 2001; Moody, 2001; Currarini et al., 2009). We explore the role of homophily in friendship formation and persistence by interacting the number of years that students i and j were randomly assigned to the same classroom (from Kindergarten to third grade for the formation analysis and from fourth grade to sixth grade for the persistence analysis) with an indicator for whether they share observable characteristics. Therefore, we estimate the following dyadic link formation model:

$$y_{ijs} = \alpha + \sum_{k=0}^3 \beta_k 1\{Years_{ijs} = k\} + \gamma x_{ijs} + \sum_{k=0}^3 \phi_k 1\{Years_{ijs} = k\} \cdot x_{ijs} + \delta_s + \mu_{ijs} \quad (2)$$

where $1\{Years_{ijs} = k\}$ denotes a binary indicator equal to 1 if students i and j were randomly assigned to the same classroom for exactly k years between Kindergarten and third grade for the formation analysis, or between fourth and sixth grades for the persistence analysis. x_{ijs} is a binary indicator equal to 1 if the link between students i and j is homophilous (i.e., $x_i = x_j$). Using the estimated parameters of interest β_k , γ , and ϕ_k we recover how the length of exposure and homophily jointly affect the probability that students i and j have a directed link in third grade and the probability that an existing link persists to sixth grade.

Identification

The empirical strategy exploits variation in shared classroom exposure across student pairs i and j within the same school. The key identification assumption is that, conditional on school fixed effects, potential friendship links between student pairs with different classroom assignments do not differ systematically on other dimensions. The school fixed effects account for school-level characteristics that are constant across students and may affect the probability of forming a directed link or a reciprocal friendship. As a result, any systematic difference in the probability of forming a directed link or a reciprocal friendship between students assigned to the same classroom and those assigned to different classrooms must be attributable to the classroom assignment itself, rather than other pre-existing differences. Importantly, this key assumption is supported by the random assignment of students to classrooms at the beginning of each grade.

A potential concern is whether the Stable Unit Treatment Value Assumption (SUTVA) holds. In our context, the variation in exposure depends entirely on random assignment. However, students may differ in their ability to make and maintain friendships, and they may interact with their classmates outside the classroom (e.g., during breaks or after school). In this case, our estimates would be a lower bound because out-of-class

interaction would produce attenuation bias. Specifically, directed links and reciprocal friendships among students assigned to different classrooms would increase, and the difference in the probability of forming a friendship, or its persistence, between that group and students assigned to the same classroom would be smaller. Reassuringly, fewer than 4% of third-grade directed and reciprocal friendships occur across classrooms as shown earlier, so any spillovers are likely limited.

The main specification in Equation 1 does not explicitly model the choice process by which students nominate their friends. Recognizing these concerns, we estimate three alternative specifications. First, we estimate a logit model to account for the binary outcome. Second, we include a student i fixed effect, which absorbs time-invariant differences in sociability, ability to make friends, and other traits that affect the propensity to nominate peers. This fixed effect also partially accounts for the heterogeneity in the decision-making process whereby students select their friends from the set of students in the school. We do not treat this specification as our preferred one because a student's sociability (captured by the i fixed effect) may itself be shaped by prior random classroom assignments, so it is potentially post-treatment and may attenuate the exposure effects. Finally, we include both student i and student j fixed effects, which separately absorb unobserved heterogeneity in sending and receiving propensities (sociability and popularity), rather than dyad-specific characteristics. Across all three alternatives, the results are similar to those from the main specification.

Finally, we estimate non-parametric dyadic-robust standard errors because errors are likely correlated across dyads that share a student, which would bias standard errors downward when clustering only at the school level (Cameron and Miller, 2014; Aronow et al., 2017; Graham, 2020). One potential solution is two-way clustering by student i and student j , but this does not fully account for the correlation across dyads. For this reason, we estimate the non-parametric dyadic-robust variance estimator proposed by

Aronow et al. (2017). The inference is robust to using this estimator.

1.3.2 Network Effects

The goal is to estimate the impact of the number of links (directed, reciprocal, followers and rejecters) a student has on her subsequent learning and non-cognitive outcomes in elementary school. The dataset allows us to construct measures of the number of friends of each type that a given student has at the school and classroom levels in each grade from fourth to sixth, based on the third-grade friendship network. Therefore, to estimate the effect of a student's network on cognitive and non-cognitive skills, we estimate the following model pooling observations from those grades for the full sample, and separately for male and female students:

$$y_{icst} = \alpha + \beta FC_{it} + \mu FS_{it} + \delta_{st} + \varepsilon_{icst} \quad (3)$$

where y_{icst} is students i 's performance (e.g., test score) in classroom c in school s at the end of grade t , where $t \in 4, 5, 6$ grades. FC_{it} and FS_{it} denote the number of friends of a given type in student i 's classroom and school, respectively, in grade t . δ_{st} is a school-by-grade fixed effects. We cluster standard errors at the school-by-grade level, which allows students' outcomes in different classrooms to be correlated within schools in a given grade. The parameter of interest, β , measures the impact of the number of classroom friends of each type on students' cognitive and non-cognitive skills.

Identification

The empirical strategy exploits variation in exposure to friends across classrooms within the same school and grade. The key identification assumption is that, conditional on school-by-grade fixed effects and student's total number of friends in the school,

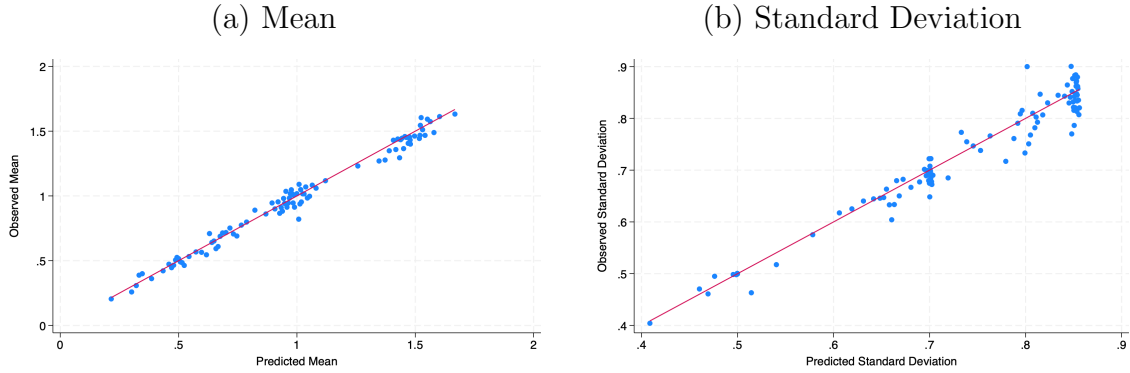
classrooms with and without a student's friend do not differ systematically along other dimensions. The school-by-grade fixed effects account for differences across cohorts within a given school and for school-grade characteristics that are constant across students. Importantly, this key assumption is supported by the random assignment of students to classrooms at the beginning of each grade.

A potential concern is weak variation in group composition, similar to challenges encountered when estimating peer effects under random assignment to classrooms (Angrist, 2014). While our paper does not rely on a linear-in-means model, the probability that a student is assigned to a classroom with none of her friends should decrease as class size increases. This issue is particularly relevant in settings with large peer groups, as they are more prone to experiencing weak variation. In our context, the variation in the number of friends in the classroom tends to decrease with class size, and identification relies on residual variation due to finite-sample imbalances in classroom assignments. However, this concern is mitigated in our setting, as most classrooms have fewer than 45 students, with an average class size of 37.

However, recognizing these potential concerns, we conduct two diagnostic analyses to assess the extent of variation in the number of friends in the classroom and to test for random assignment. First, we assess whether the number of friends in the classroom that a student has follows a hypergeometric distribution with school- and student-specific parameters. Panel A in Figure 1.3 shows the observed mean and the predicted mean based on the hypergeometric distribution. On average, students get exactly as many friends (0.98) in the classroom as the hypergeometric predicts (0.99), indicating no systematic sorting of friends into classrooms. Panel B of the same figure analyzes the dispersion of the number of friends. The empirical dispersion closely tracks the distribution's predicted spread, and only 2.01% of classrooms fall outside the 95% interval implied by the hypergeometric calculation. Taken together, these diagnostics indicate that the

observed variation in classroom friend counts is what we would expect under random assignment.

Figure 1.3: Observed Vs Predicted Mean and Standard Deviations of Friends in Classroom



Notes: Panel A in the figure above shows the observed mean of the number of directed friends in the classroom and the expected mean coming from the hypergeometric distribution based on the school- and student-specific parameters. Panel B in the figure shows a similar analysis but using the observed and expected standard deviations. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates.

Second, we test whether the observed probability of having k friends in the classroom, conditional on the number of friends in the school, matches the expected probability implied by the hypergeometric distribution. We simulate the probability of having k friends in the classroom based on the school- and student-specific parameters. Conditioning on the number of friends in the school holds constant degree heterogeneity, so differences in the probabilities of having k friends in the classroom are not driven by some students having more (or fewer) friends overall. Appendix Figure A.1 reports the observed and expected probabilities separately for students who have 1, 2 or 3 friends in the school. The figure shows that the observed and expected probabilities match for each number of friends in the classroom, conditional on the number of friends in the school.

Finally, the ability to exploit friends composition relies on sufficient residual variation in the key variables after accounting for fixed effects and the total number of friends in the school. Given the random assignment, the distribution of the number of friends in the classroom, conditional on school-by-grade fixed effects and the number of friends in school, should be approximately normal. Appendix Figure A.2 shows that the distribution of the residualized number of directed links and reciprocal friends in the classroom is indeed approximately normal, consistent with random assignment. More importantly, the figure also shows some within school-grade variation in the number of friends in the classroom, further supporting the validity of our empirical strategy.

Additionally, to further support the causal interpretation of the estimates, we provide several sensitivity tests. First, we conduct sensitivity tests on the functional form, controls, and sample definition by estimating three alternative specifications: i) non-linear effects in the number of friends in the school to allow for a non-parametric estimation; ii) we include teacher controls to account for the possibility that teachers may influence the classroom environment and affect student outcomes; and (iii) we restrict the sample to fourth grade only, the year immediately following the friendship data collection which minimizes network mismeasurement and reduces attrition concerns. As we show later, the results from these specifications are similar to those obtained using the main specification.

Second, we test the sensitivity of the estimates to the exclusion of one school-grade at a time to ensure that the results are not driven by a single school in a given grade. Appendix Figure A.3 shows the distribution of the leave-one-out coefficient estimates. In particular, one unique regression 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 centers around the value of the coefficient estimate found in the main analysis.

1.4 Results

1.4.1 Friendship Formation and Persistence

We estimate the regression model specified in Equation 1 for directed links and reciprocal friendships using all possible dyads of students (i, j) . Columns (1) and (3) in Table 1.3 show that being randomly assigned to the same classroom increases the probability that a pair of students has a directed link or a reciprocal friendship. For instance, being randomly assigned to the same classroom in second grade increases the probability of having a directed link by 1.2 percentage points, which corresponds to a 41% increase relative to the mean.²³ For reciprocal friendships the effect corresponds to a 60% increase relative to the mean. Interestingly, the magnitude of the effects increases with grade level. This result is not surprising as children at this age spend most of the day in the same classroom with the same group of peers and teachers. As a result, the set of people they can interact with is restricted to their peers in the classroom, and this increased interaction makes them more likely to form friendships which is consistent with our descriptive evidence discussed above.

Furthermore, the length of exposure also plays a significant role in friendship formation. In particular, being randomly assigned to the same classroom for multiple years fosters repeated interactions and increases the likelihood that students i and j meet through mutual friends. Columns (2) and (4) in Table 1.3 show that each additional shared year increases the probability of having a directed link by 2.2 percentage points (a 76% increase relative to the mean). Regarding reciprocal friendships, the effect corresponds to an 80% increase relative to the mean. Standard errors are smaller for reciprocal friendships, which is consistent with the lower variance of the outcome and with reciprocal ties likely being measured with less noise because they represent stronger ties.

²³The mean in this table is small because it includes all possible student dyads (i, j) .

Table 1.3: Impacts on Friendship Formation

	Directed		Reciprocal	
	(1)	(2)	(3)	(4)
Same classroom in K	0.003*** (0.001)		0.001*** (0.000)	
Same classroom in 1st	0.005*** (0.001)		0.003*** (0.000)	
Same classroom in 2nd	0.012*** (0.001)		0.006*** (0.000)	
Same classroom in 3rd	0.071*** (0.002)		0.024*** (0.001)	
Years in same classroom		0.022*** (0.001)		0.008*** (0.000)
Mean of Dependent Variable	0.029	0.029	0.010	0.010
Observations	707568	707568	353784	353784
School FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions on a binary indicator of whether students i and j have a directed link or are reciprocal friends in third grade. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other. Columns (1) and (3) present the results using binary indicators for whether the students were randomly assigned to the same classroom on each grade between Kindergarten and 3rd grade. Columns (2) and (4) present the results using the total number of years the students were randomly assigned to the same classroom between Kindergarten and 3rd grade. All regressions are limited to schools with at least two classrooms per grade. *** 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.

A novel aspect of our study is that we can analyze friendship persistence patterns given

longitudinal data on friendship links. For this reason, we estimate Equation 1 separately for pairs of students who had a directed link in third grade and pairs who did not have a directed link. Table 1.4 shows that being randomly assigned to the same classroom increases the probability that a pair of students have a directed link in sixth grade for both groups. However, when comparing the estimates across groups, we observe that, for those who were already friends, the effects are larger and statistically different from those for pairs that were not previously friends. For instance, being randomly assigned to the same classroom in sixth grade increases the probability of a directed link in sixth grade by 20.8 percentage points for those who were friends in third grade and by 6.7 percentage points for those who were not previously friends. As in the formation analysis, each additional year of exposure also increases the probability that the friendship persists.²⁴

One concern about the results is whether the directed links provide a meaningful tie comparable to the stronger ties reflected in reciprocal friendships. While we cannot conclusively estimate the strength of these directed ties, we provide two pieces of evidence. First, based on the persistence results, if directed links in third grade were not meaningful, we would expect them not to persist. However, as discussed earlier, the effects on friendship persistence are larger and statistically different for pairs who were already friends in third grade than for those who were not. Second, if the links were not meaningful, we would expect zero exposure between fourth and sixth grades to have no impact on persistence. Instead, Appendix Figure A.4 shows that among pairs who were friends in third grade, having zero shared classrooms between fourth and sixth grades is still associated with a roughly 16% probability of persisting, whereas among pairs who were not friends, the corresponding probability is near zero. Taken together, these patterns suggest that directed links were sufficiently meaningful to persist and exhibit some degree of strength.

²⁴Appendix Table A.1 shows the results for reciprocal friendships. Overall, the results are similar to those found for directed links but the magnitudes in percentage points are smaller.

Table 1.4: Impacts on Persistence of Directed Friendships

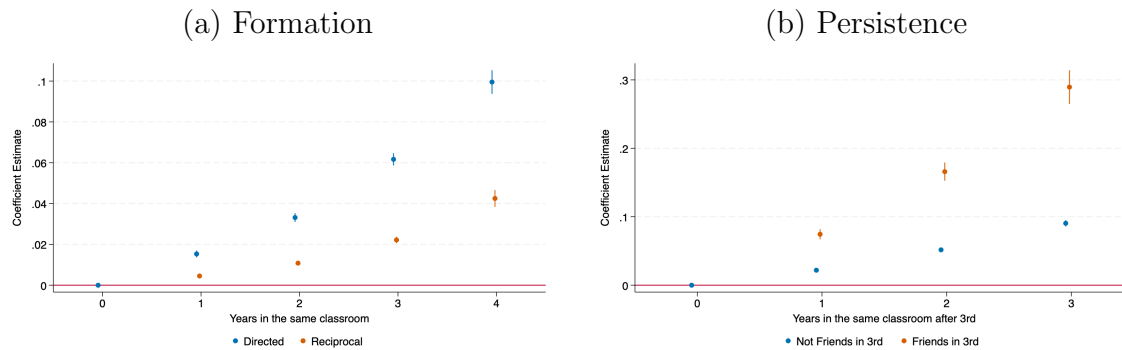
	Not Friends in 3rd		Friends in 3rd	
	(1)	(2)	(3)	(4)
Same classroom in 4th	0.005*** (0.000)		0.026*** (0.004)	
Same classroom in 5th	0.012*** (0.001)		0.037*** (0.004)	
Same classroom in 6th	0.067*** (0.002)		0.208*** (0.006)	
Years in same classroom after 3rd		0.028*** (0.001)		0.091*** (0.003)
Mean of Dependent Variable	0.029	0.029	0.119	0.119
Observations	990641	990641	27593	27593
School FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions on a binary indicator of whether students i and j have a directed link in sixth grade, separately for pairs of students who were friends in third grade and pairs that were not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. Columns (1) and (3) present the results using binary indicators for whether the students were randomly assigned to the same classroom on each grade between 4th and 6th grade. Columns (2) and (4) present the results using the total number of years the students were randomly assigned to the same classroom between 4th and 6th grade. All regressions are limited to schools with at least two classrooms per grade. *** 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.

Nevertheless, the previous results on the length of exposure assume that the effect

of each additional year is linear. Figure 1.4 reports estimates of the nonlinear effects on friendship formation separately for directed links and reciprocal friendships and on persistence of directed links separately for those who were friends in third grade and those who were not. Panel A presents the results for formation and indicates that the marginal impact of an extra year of exposure is nonlinear. Specifically, the magnitude of the effect increases with the number of years spent together in the same classroom for both directed links and reciprocal friendships. For example, being in the same classroom for two years increases the probability of having a directed link by 3.3 percentage points, while being together for four years increases it by 9.9 percentage points. Similarly, Panel B presents the persistence results. As before, the impact increases with the number of years spent together, and the effects are larger for those who were friends in third grade. Overall, these results suggest that repeated exposure and interaction over multiple years significantly enhance the formation and persistence of friendship links.

Figure 1.4: Impacts on Friendship Formation and Persistence by Years of Exposure



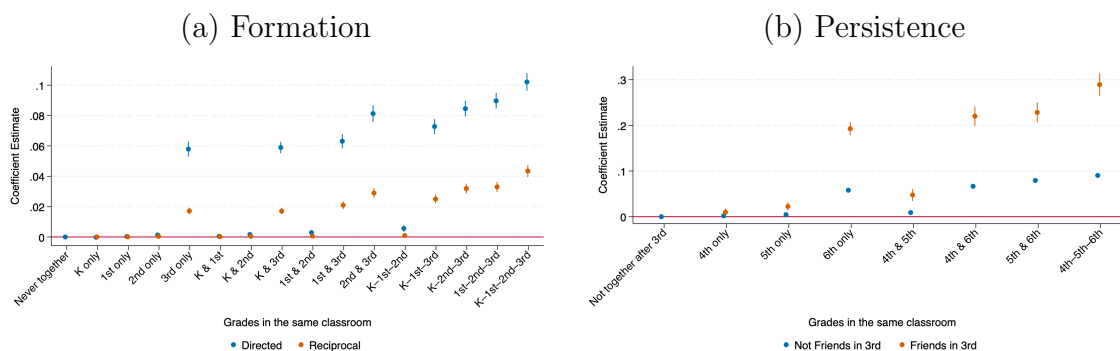
Notes: The figure above reports non-linear estimates from regressions of binary indicators for the number of years of exposure. Panel A presents the results for friendship formation for both directed and reciprocal links. Panel B presents the results for friendship persistence only for directed links, separately for those who were friends in third grade and those who were not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other. 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.

The results on the length and timing of exposure suggest that their interaction could also play a role on the formation and persistence of friendships. Figure 1.5 reports estimates of the interaction between exposure length and grade timing, using pairs who were never in the same classroom as the comparison group. In Panel A, we present the effects on friendship formation for directed and reciprocal friendships. For a fixed number of shared years, effects are larger when exposure occurs in the most recent grades, suggesting that third grade has a particularly strong effect, as discussed before. For instance, among pairs with three shared years, the effect of being randomly assigned together in first, second, and third grades is larger than the other three-year sequences, while the effect of being randomly assigned in Kindergarten, first, and second is the

smallest. In Panel B, we present the effects on friendship persistence, and a similar pattern emerges.

For persistence, we separate pairs by whether they were friends in third grade to distinguish continuation of existing ties from the creation of new ones. We find that being assigned to the same classroom in sixth grade increases the probability of a directed link by 20.8 percentage points for those who were friends in third grade and by 6.7 percentage points for those who were not previously friends.

Figure 1.5: Impacts on Friendship Formation and Persistence by Timing and Years of Exposure

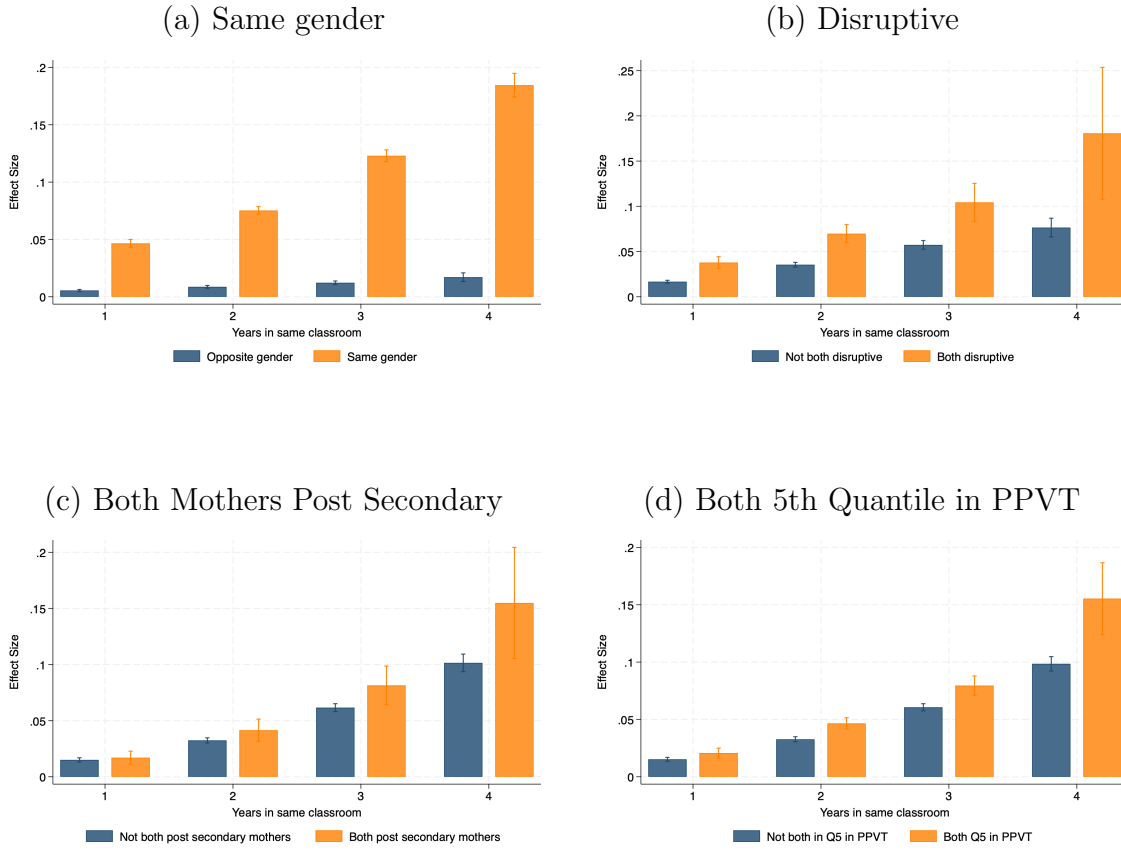


Notes: The figure above reports estimates from regressions of binary indicators for the sequences of the interaction between the timing and the length of exposure. Panel A presents the results for friendship formation for both directed and reciprocal links. Panel B presents the results for friendship persistence only for directed links, separately for those who were friends in third grade and those who were not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other. 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.

Role of Homophily

Given that the descriptive evidence presented above suggests some degree of homophily in the formation of friendships, we estimate the regression model specified in Equation 2 and use the estimates to calculate the predicted increase in the probability of having a directed link for eight different groups. These groups are defined by the interaction between the number of years a given pair of students was randomly assigned to the same classroom and whether they share the same characteristic. We conduct this analysis separately for four characteristics: (i) same-gender pairs compared with opposite-gender pairs; (ii) pairs of students that are both disruptive compared with pairs where one is disruptive and one is not; (iii) pairs of students that both have post-secondary educated mothers compared with pairs where one does and one does not; and (iv) pairs of students where both are in the highest quintile of the PPVT distribution compared with pairs where one is and one is not. It is important to note that for disruptiveness, post-secondary maternal education, and PPVT, we exclude pairs of students in which neither student has the focal characteristic. For instance, in the disruptive analysis, pairs in which both students are not disruptive are excluded.

Figure 1.6: Impacts on Friendship Formation by Years of Exposure and Homophily



Notes: The figure above reports estimates of the predicted increase of the probability that students i and j have a directed link. The estimates were obtained from regressions of binary indicators of the number of years students i and j are randomly assigned to the same classroom interacted with a binary indicator of homophily among four characteristics: (i) same-gender pairs compared with opposite-gender pairs; (ii) pairs of students that are both disruptive compared with pairs where one is disruptive and one is not; (iii) pairs of students that both have post-secondary educated mothers compared with pairs where one does and one does not; and (iv) pairs of students where both are in the highest quintile of the PPVT distribution compared with pairs where one is and one is not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. All regressions are limited to schools with at least two classrooms per grade. Vertical lines represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Figure 1.6 presents estimates of the predicted increase in the probability that students i and j have a directed link obtained by interacting indicators for the number of shared classroom years with an indicator for homophily along the characteristics discussed above.. It shows that the probability of having a directed link increases with the number of years spent together across all groups and characteristics. Panel A shows a remarkable difference between same-gender and opposite-gender pairs. In particular, among pairs randomly assigned together for three years, same-gender pairs experience larger increases (around 0.12 percentage points) in friendship formation than opposite-gender pairs (around 0.02 percentage points). These results are consistent with previous findings that, during elementary school, female students tend to befriend other females and male students tend to befriend other males (McPherson et al., 2001; Garrote et al., 2023). Panels C and D report results for maternal education and PPVT and likewise indicate that homophilous links are more likely to form.

Interestingly, Panel B shows that pairs in which both students are disruptive are more likely to become friends than those in which one student is and one is not. However, this result could be confounded by the gender composition of disruptive students, as 18% of male students are identified as disruptive, while less than 3% of female students are. For this reason, in Appendix Figure A.5, we analyze the effects of disruptiveness separately for pairs of students who are both male (Panel A) and pairs of students who are both female (Panel B). As expected, the Figure shows that the results are driven by the male pairs and hold when examining within male pairs.

Finally, Appendix Figure A.6 presents the results of the predicted increase in the probability of a directed link persists for eight different groups using the same four characteristics. Overall, the Figure shows a similar pattern to that observed in the formation figure. However, the differences between homophilous pairs and the non-homophilous pairs appear to be smaller than those observed in formation. This pattern is

consistent with the fact that, as students grow, they start interacting with more peers and making new friends, which helps explain the smaller differences.

1.4.2 Network Effects

We estimate the regression model specified in Equation 3 separately for each cognitive and non-cognitive test score. Table 1.5 shows that having more friends in the classroom reduces test scores on math and executive function tests. For instance, having one additional directed friend in the classroom reduces math test scores by 0.006 standard deviations. Similarly, one additional directed friend in the classroom is associated with a 0.023 standard deviation decrease in executive function. Interestingly, the effect is driven by the number of rejecters in the classroom, which could reflect a behavioral response, as students are negatively affected when they are randomly assigned to a classroom with peers they believe are friends, but who do not reciprocate.

Table 1.5: Impact of Friendships on Test Scores

	Directed	Reciprocal	Followers	Rejecters
	(1)	(2)	(3)	(4)
A. Math				
No. of Friends in Classroom	-0.006*	-0.003	0.002	-0.007*
	(0.003)	(0.006)	(0.004)	(0.004)
Mean of Dependent Variable	0.031	0.031	0.031	0.031
Treatment Effect of 1SD increase	-0.005	-0.002	0.002	-0.005
Observations	41431	41431	41431	41431
B. Executive Function				
No. of Friends in Classroom	-0.023**	0.001	0.001	-0.031***
	(0.009)	(0.014)	(0.010)	(0.011)
Mean of Dependent Variable	0.034	0.034	0.034	0.034
Treatment Effect of 1SD increase	-0.019	0.000	0.001	-0.024
Observations	14707	14707	14707	14707
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the number of friends in the classroom on math scores, conditional on the school-by-grade fixed effects and the number of friends in the school. Each column represents one of the friendship types analyzed: directed links (i.e., student i nominates student j , regardless of whether student j reciprocates), reciprocal friendships (i.e., both student i and student j nominate each other), followers (i.e., those who nominated a student but were not reciprocally nominated) and rejecters (i.e., those who are nominated by a classmate but do not reciprocate). 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-by-grade level.

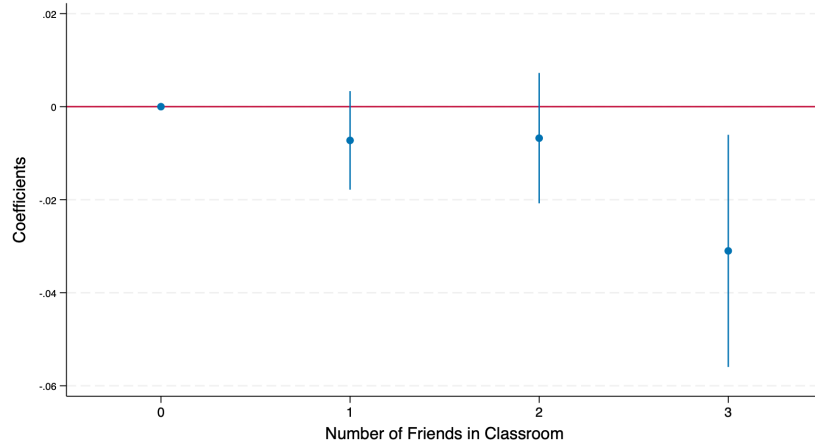
There are at least three plausible explanations for our findings. First, the presence of directed friends, particularly rejecters, may lower students' happiness, reducing effort and

self-confidence, which can negatively affect academic performance as they could be trying to get attention. Second, the characteristics of friends could negatively affect academic results. For instance, in the previous section, we show that disruptive students tend to befriend one another, which could depress test scores, in line with previous evidence on the effects of neglected and abused students (Santavirta and Sarzosa, 2025). Finally, teachers may adjust their teaching strategies and classroom management, potentially mediating the effects on learning outcomes.

Given the potential differential effects by gender, Table 1.7 shows that male students are negatively affected by the presence of directed links or rejecters in the classroom, while female students are not. These results highlight the mixed evidence on whether friendship networks improve performance and how the effects differ by gender. On the one hand, working with friends could improve performance if it increases motivation or social pressure (Babcock et al., 2015). On the other hand, friendship networks may reduce performance if they are used for socializing or if in groups with less able friends (Bandiera et al., 2010). Therefore, in Section 1.4.3, we analyze potential mechanisms that could drive these results and explain the observed gender differences.

Additionally, we estimate the non-linear effects of the number of friends in the classroom. Accordingly, we create binary indicators of whether a student has zero, one, two, or three friends in the classroom. Figure 1.7 shows that negative effects on test scores emerge only when a student has three friends in the classroom. This result suggests that as the size of the friendship network increases, potentially increasing opportunities to socialize, negative effects may appear.

Figure 1.7: Non-Linear Effects of Friendship Links on Math Scores



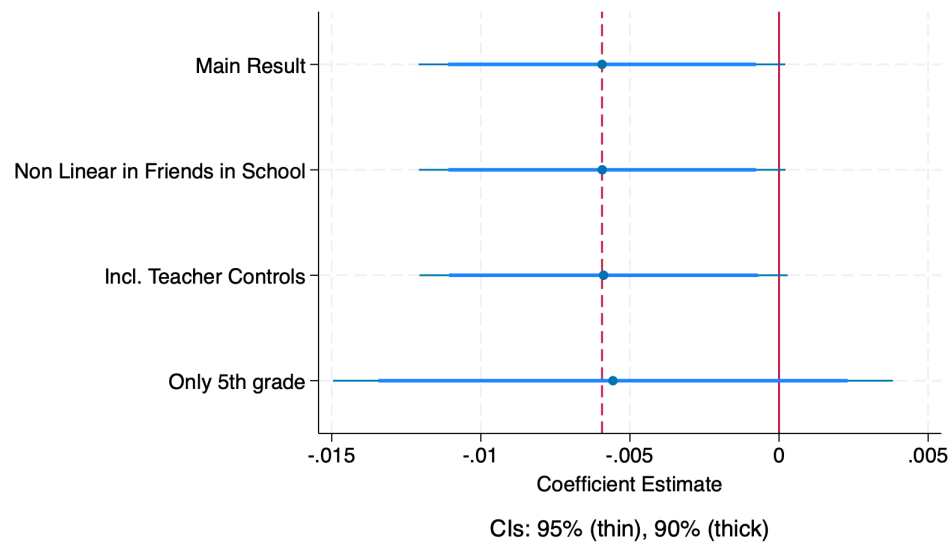
Notes: The figure above reports non-linear estimates from regressions of binary indicators for the number of directed friends in the classroom on math scores, conditional on the school-by-grade fixed effects and the number of directed friends in school. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates, and a reciprocal friendship is defined as one in which both student i and student j nominate each other. 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. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Robustness checks and Alternative Specifications

This section provides a brief summary of three alternative specifications and robustness checks to address concerns about the main specification. The first check estimates the non-linear effects for the number of friends in the school to allow for a non-parametric estimation and is motivated by evidence that the impact of classroom friend counts is non-linear. The second includes teacher controls to account for how teachers may influence the classroom environment and affect student outcomes through their interactions and management. Finally, we restrict the sample to fourth grade only, the year immediately

following the friendship data collection and the grade least affected by attrition.

Figure 1.8: Robustness Checks on Math Scores



Notes: The figure above reports estimates from regressions of the number of directed friends in the classroom on math scores, conditional on the school-by-grade fixed effects and the number of directed friends in school. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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. Wider horizontal bars represent 95% confidence intervals and darker horizontal bars represent 90% 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 in Table 1.5. All other estimates are variations on the baseline model. Estimate 2 includes the number of friends in a non-linear way. Estimate 3 includes teacher controls, and estimate 4 restricts the sample to fourth grade only.

Figure 1.8 presents the results of these alternative specification checks. Following the main result (i.e., the main specification), the next three estimates show that the results are

insensitive to these robustness checks. The wider confidence intervals represent the 95% confidence intervals and the darker ones represent the 90% confidence intervals. Overall, the first two robustness checks are similar in magnitude and significance to the main result, while the last one is similar in magnitude but not statistically significant. However, the estimates using the restricted sample should be interpreted cautiously, as the smaller sample size leads to noisier estimates, and our power to detect effects is reduced.

1.4.3 Potential Mechanisms of the Network Effects

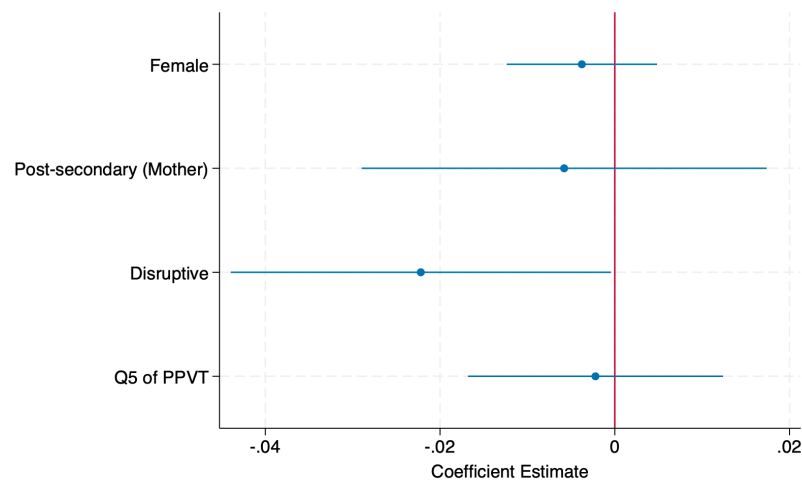
Having shown that the number of friends in the classroom, particularly directed and rejecters, negatively affects test scores in both cognitive (math and executive function) skills, we now explore potential mechanisms that might explain these results. We focus on three main channels: (i) the role of friends' characteristics, (ii) the effects on mental health and non-cognitive skills, and (iii) the role of teacher quality.

Friend Characteristics

In this section, we explore the role of the friends' characteristics in mediating the negative effects of friendship networks on cognitive skills. First, we construct the number of directed friends at the classroom and school levels that a student has who are female, whose mothers have post-secondary education, who are disruptive, or who are in the highest quintile of the PPVT distribution. Using this information, we estimate the regression model specified in Equation 3 separately for each friend characteristic. Figure 1.9 presents the results of these regressions. In this figure, each line plots regression estimates for math test scores using the numbers of friends at the classroom and school levels with the given characteristic. This figure shows that the numbers of female friends, friends with post-secondary educated mothers, and friends in the highest quintile of the

PPVT distribution are not associated with lower math test scores. However, the number of disruptive directed friends in the classroom has a negative effect on math scores. These findings align with prior literature showing that exposure to disruptive peers negatively affects test scores (Santavirta and Sarzosa, 2025).

Figure 1.9: Effect of Directed Friendships on Math Scores by Friend Characteristics



Notes: The figure above reports estimates from regressions of the number of directed friends that have a certain characteristic in the classroom on math scores, conditional on the school-by-grade fixed effects and the number of directed friends that have a certain characteristic in school. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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. Each line plots regression estimates for math test scores using the number of friends at the classroom and school levels with the given characteristic.

Table 1.6: Impact of Friend Characteristics of Directed Friendships on Math Test Scores by Gender

	Female	Post-secondary	Disruptive	PPVT
	(1)	(2)	(3)	(4)
A. Male				
No. of Friends in Classroom	-0.002 (0.012)	-0.018 (0.017)	-0.022* (0.013)	-0.003 (0.011)
Mean of Dependent Variable	0.091	0.091	0.091	0.091
Treatment Effect of 1SD increase	-0.001	-0.004	-0.007	-0.001
Observations	21308	21308	21308	21308
B. Female				
No. of Friends in Classroom	-0.005 (0.005)	0.007 (0.016)	-0.030 (0.028)	-0.001 (0.011)
Mean of Dependent Variable	-0.032	-0.032	-0.032	-0.032
Treatment Effect of 1SD increase	-0.004	0.002	-0.004	-0.000
Observations	20123	20123	20123	20123
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the number of friends that have a certain characteristic in the classroom by gender, conditional on the school-by-grade fixed effects and the number of friends that have a certain characteristic in the school. Each column represents one of the friends' characteristics analyzed: female, maternal education, disruptiveness and prior PPVT scores. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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-by-grade level.

Given the previous results on how homophily plays a role in the formation and persistence of friendships and how the negative effects on cognitive scores were experienced

only by male students, we separate this analysis by gender. Table 1.6 shows the results of these regressions. In this table, each column reports regression estimates for math test scores using the numbers of friends at the classroom and school levels with the given characteristic. This table shows no negative effects for female students, whereas for male students, we find negative effects of the number of disruptive directed friends in the classroom. These findings, combined with the evidence that male students are more likely to be disruptive, suggest that when disruptive students are together in the classroom, they may not behave properly and thereby affect learning.

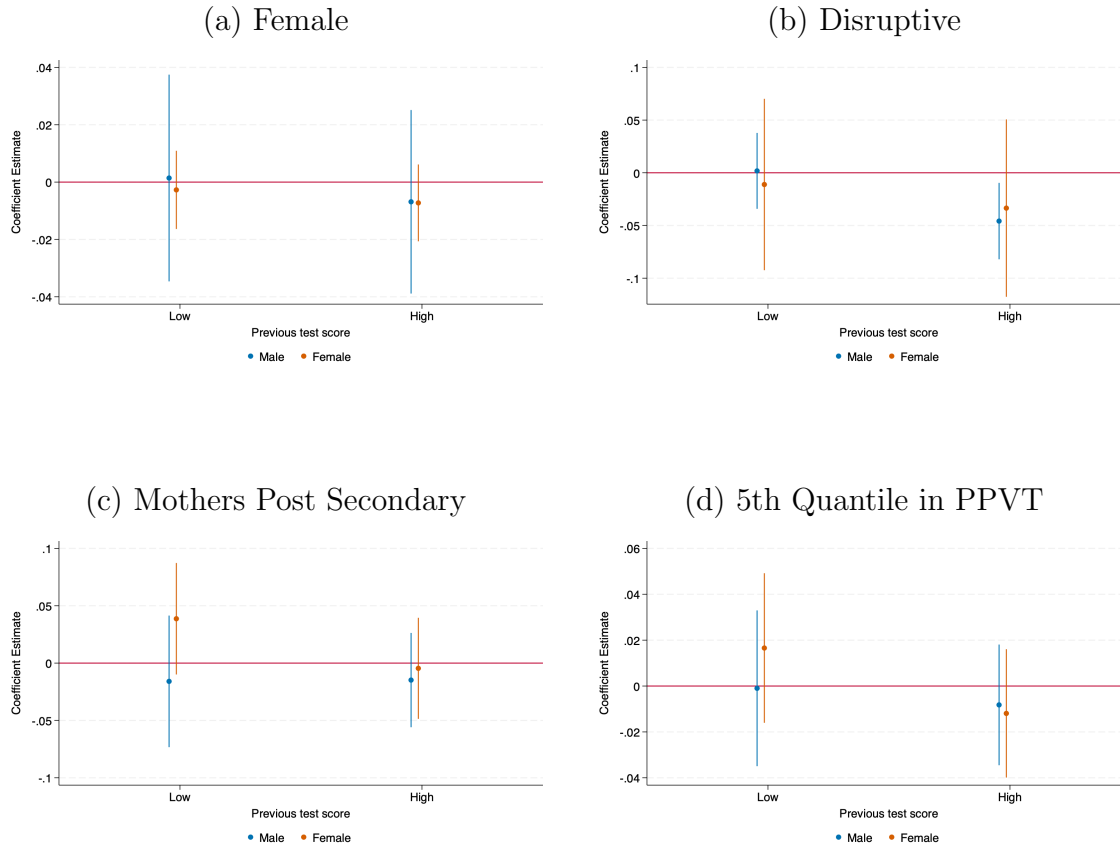
Finally, we analyze how ability plays a role in the results, given that previous evidence finds that when students are placed together with less able friends, the effects of friendship networks are negative (Bandiera et al., 2010). Figure 1.10 divides the sample into two groups based on prior math scores (i.e., above and below the median) and presents regression results for each group by gender. Overall, there do not appear to be differential effects by ability. However, for high-ability male students, we find negative effects. These findings align with previous evidence, as disruptive students in our sample tend to have lower math test scores than non-disruptive students. Therefore, when high-ability male students are randomly assigned to classrooms with less able friends, they experience reductions in test scores.

Table 1.7: Impact of Friends on Math Test Scores by Gender

	Directed	Reciprocal	Followers	Rejecters
	(1)	(2)	(3)	(4)
A. Male				
No. of Friends in Classroom	-0.009** (0.004)	-0.003 (0.008)	0.008 (0.006)	-0.011** (0.005)
Mean of Dependent Variable	0.091	0.091	0.091	0.091
Treatment Effect of 1SD increase	-0.008	-0.002	0.008	-0.008
Observations	21308	21308	21308	21308
B. Female				
No. of Friends in Classroom	-0.003 (0.005)	-0.005 (0.008)	-0.003 (0.005)	-0.002 (0.006)
Mean of Dependent Variable	-0.032	-0.032	-0.032	-0.032
Treatment Effect of 1SD increase	-0.002	-0.003	-0.003	-0.001
Observations	20123	20123	20123	20123
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the number of friends in the classroom on math scores by gender, conditional on the school-by-grade fixed effects and the number of friends in the school. Each column represents one of the friendship types analyzed: directed links (i.e., student i nominates student j , regardless of whether student j reciprocates), reciprocal friendships (i.e., both student i and student j nominate each other), followers (i.e., those who nominated a student but were not reciprocally nominated) and rejecters (i.e., those who are nominated by a classmate but do not reciprocate). 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-by-grade level.

Figure 1.10: Effect of Directed Friendships on Math Scores by Gender, Friend Characteristics, and Previous Test Results



Notes: The figure above reports estimates from regressions of the number of directed friends that have a certain characteristic in the classroom on math scores, conditional on the school-by-grade fixed effects and the number of directed friends that have a certain characteristic in school. Each panel shows the results for a certain characteristic by gender and the test result in the previous year. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school-by-grade level.

Effects on other outcomes

In the previous section, we show that the friendship network effects appear to be mediated by homophily. In this section, we explore how friendship networks affect other outcomes, particularly mental health and non-cognitive skills. First, Table 1.8 shows reductions in depression associated with the presence of both directed and reciprocal friendships, suggesting that, if anything, the presence of friends in the classroom creates a positive spillover on students' mental health. This rules out the possibility that the effects are driven by lower happiness or self-confidence. Moreover, this result, combined with the results on disruptive friends, suggests that negative effects on test scores likely operate through a distraction or lack of attention channel. Finally, Appendix Table A.2 shows no significant effects on the other non-cognitive skills (growth mindset and grit).

Teacher Quality

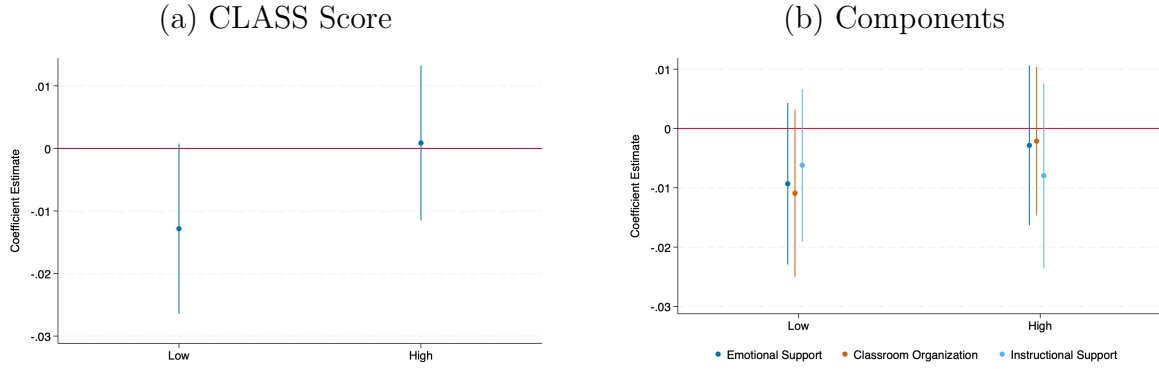
Given the existing literature on teacher quality and its positive effect on test scores, we test whether teacher behaviors, as measured by the CLASS, can mitigate the negative friendship network effects. Panel A in Figure 1.11 shows that high-quality (those above the median) teachers reduce the negative effects. However, when separating the results by CLASS domain, Panel B does not show any differences between high-quality and low-quality teachers. While we cannot conclusively identify what high-quality teachers do to mitigate the effects, these findings suggest that high-quality teachers are better at managing the negative externality, particularly those stemming from disruptive friends. More importantly, this highlights the role of classroom management and the importance of appropriate incentives to improve teachers' performance that could positively affect student outcomes.

Table 1.8: Impact of Friendships on Non-Cognitive Skills

	Directed	Reciprocal	Followers	Rejecters
	(1)	(2)	(3)	(4)
A. Depression				
No. of Friends in Classroom	-0.034** (0.015)	-0.053** (0.026)	-0.012 (0.019)	-0.021 (0.019)
Mean of Dependent Variable	-0.007	-0.007	-0.007	-0.007
Treatment Effect of 1SD increase	-0.028	-0.029	-0.011	-0.015
Observations	7535	7535	7535	7535
B. Self-esteem				
No. of Friends in Classroom	-0.017 (0.018)	-0.008 (0.025)	0.022 (0.019)	-0.020 (0.022)
Mean of Dependent Variable	-0.002	-0.002	-0.002	-0.002
Treatment Effect of 1SD increase	-0.014	-0.005	0.021	-0.014
Observations	7535	7535	7535	7535
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the number of friends in the classroom on non-cognitive skills, conditional on the school-by-grade fixed effects and the number of friends in the school. Each column represents one of the friendship types analyzed: directed links (i.e., student i nominates student j , regardless of whether student j reciprocates), reciprocal friendships (i.e., both student i and student j nominate each other), followers (i.e., those who nominated a student but were not reciprocally nominated) and rejecters (i.e., those who are nominated by a classmate but do not reciprocate). All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, and age squared. *** 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-by-grade level.

Figure 1.11: Heterogeneous Effects of Directed Friendships on Math Scores by Teacher Quality



Notes: The figure above reports estimates from regressions of the number of directed friends in the classroom on math scores by teacher quality measured using the CLASS (*Classroom Assessment Scoring System*) score, conditional on the school-by-grade fixed effects and the number of directed friends in school. Teachers are separated into two groups: low and high quality, using the median score as the cutoff. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

1.5 Conclusion

This paper analyzes the impact of timing and length of exposure on friendship formation and persistence, and the impact of friendship networks on students’ cognitive and non-cognitive skills in elementary school setting. These data come from a unique longitudinal experiment in Ecuador in which children are randomly assigned to classrooms at the beginning of each school year for seven consecutive years. Compliance with the random assignment was nearly perfect, averaging 98.9% over the seven years.

The random assignment creates variation along two dimensions that we leverage. First, the variation in the timing and length of exposure between two students within a school is random. Notably, in our data, two students with the same underlying ability who attend the same school can experience different exposure to their peers because they are randomly assigned to different classrooms, each with slightly different peers. Therefore, we compare pairs of children who were randomly assigned to the same classroom with pairs who were not. Second, conditional on school-by-grade fixed effects and the student's number of friends in the school, the variation in the number of friends a student has in the classroom is random. Therefore, in our data, two students with the same underlying ability who attend the same school and have the same number of friends in that school can be placed together in classrooms with different numbers of friends due to the random assignment.

First, we find that both the timing and length of exposure positively affect the formation and persistence of friendship links. Additionally, we document the extent of homophily in these friendship links across four characteristics (gender, disruptiveness, maternal education, and PPVT scores) and show that homophilous links are more likely to form and persist. Second, we find that the number of friends in the classroom negatively affects math and executive function scores, but we find reductions in depression, suggesting that the results are not driven by lower happiness or self-confidence. Moreover, we show that these negative effects are concentrated among male students and are driven by the number of disruptive directed friends in the classroom, suggesting that the effects likely operate through a distraction or lack of attention channel. These findings contribute to the understanding of how friends' characteristics and gender composition influence student outcomes during formative years. Finally, having higher-quality teachers can help mitigate the adverse effects on learning, but future research should explore and identify what, specifically, those teachers do.

This research has several advantages compared to previous studies. First, we leverage the random assignment in elementary schools to analyze how exposure affects friendship formation and persistence in a developing country in Latin America. This setting allows us to explore the variation in the timing and length of exposure to a peer in the classroom, overcoming any selection and reflection biases present in other network effects studies. Second, we leverage the random assignment to analyze how friendship networks affect cognitive and non-cognitive skills. This allows us to show that networks have potential long-term effects as they affect the human capital accumulation process. However, we also provide evidence addressing the concerns about weak variation, similar to those encountered when estimating peer effects under random assignment to classrooms. Third, we explore the role of homophily in friendship formation and persistence, and as a potential mechanism mediating the observed network effects. Our results highlight the importance of understanding how friendships form and whether they persist, and their potential impacts on cognitive and non-cognitive skills. More importantly, we provide insights into the complex dynamics of friendship networks in early education settings.

Finally, given that network effects are context-specific and depend on the outcome being studied, these results raise essential policy questions about how to optimally allocate children to classrooms, how to address adverse effects in Latin America, and which policies are most effective, especially because these effects stem from a particular group of peers (i.e., friends). Future research should explore the underlying mechanisms through which friendship networks affect test scores, how other factors influence the network formation process and its persistence, and whether the effects persist into high school or adulthood.

Chapter 2

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

2.1 Introduction

Peer effects significantly shape behavior, performance, and decision-making across education, the workplace, and social environments.¹ In education, previous studies show that classroom composition influences student outcomes through various channels, including class size (Hoxby, 2000b), peers' academic achievement (Hanushek et al., 2003), gender composition (Hoxby, 2000a), social interactions (Card and Giuliano, 2013), and relative class rank (Murphy and Weinhardt, 2020; Carneiro et al., 2025b). Given the long-term consequences of early childhood development, it is critical to understand how peer composition influences educational outcomes in elementary schools, especially

¹See Carrell et al. (2009); Sacerdote (2011); Card and Giuliano (2013); Sacerdote (2014); Feld and Zölitz (2017, 2022) for education settings; Guryan et al. (2009); Cornelissen et al. (2017); Villeval (2020) for workplace settings.

in developing countries where resource constraints and institutional incentives differ substantially.

The impact of exposure to high-achieving peers is theoretically ambiguous and empirically mixed. Although some studies find positive effects on achievement (Balestra et al., 2023), aspirations, and long-term outcomes (de Gendre and Salamanca, 2020; Bertoni et al., 2020), others find negative impacts, particularly when exposure lowers students' relative class rank or increases competition (de Roux and Riehl, 2022; Chen and Hu, 2024).² However, most of this research focuses on older students in developed countries and emphasizes test scores as the primary outcome. In this paper, we study how high-achieving peers affect both cognitive and non-cognitive skills throughout elementary school, an early and critical period when peer exposure is intense and in which cognitive and non-cognitive skills are strong predictors of long-term outcomes (Moffitt et al., 2011; Heckman et al., 2006, 2013).

We estimate how the proportion of high achievers, identified through teacher rankings of students' performance, affects cognitive (math and executive function) and non-cognitive skills (depression, self-esteem, growth mindset, and grit). We use data from a unique randomized experiment involving the 2012 entering Kindergarten cohort across 202 elementary schools in Ecuador. For this cohort, we observe rich longitudinal data spanning seven consecutive years, including baseline characteristics, test scores, non-cognitive skills, teacher characteristics, and teacher quality indicators.

At the beginning of each grade, from Kindergarten through 6th grade, students were randomly assigned to classrooms within their schools, each of which had at least two classrooms per grade. Compliance with the random assignment was nearly perfect (98.9% on average). Students who remained in the same school thus experienced seven exogenous,

²These effects may also vary across student subgroups, such as high-achieving girls (Cools et al., 2022) or students from more advantaged backgrounds (Bertoni et al., 2020).

orthogonal peer groups. The variation in the proportion of high achievers across classrooms is attributable to this random assignment, enabling us to causally identify the effects of high achievers on both cognitive and non-cognitive skills.³

We identify high achievers using information reported by their teachers. At the end of each grade, we asked teachers to identify the five students with the highest learning in their class. A student is classified as a high achiever if at least 50% of the teachers who observed them in previous grades identified them as such, reducing concerns about measurement error or biases in any single teacher. Based on this classification, we construct the leave-one-out proportion of high achievers in each classroom.

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 exogenous variation in the proportion of high achievers. 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, a set of basic self-regulatory skills, by 0.014 SD. Alongside these reductions in cognitive skills, we find no effect on non-cognitive skills in 6th grade and a decrease in self-reported happiness in 1st grade. These results are robust across various specifications, estimation methods, and alternative definitions of high achievers (detailed in section 2.4.2). Moreover, when separating by gender, we find that male high achievers affect only male students, and female high achievers affect only female students. This pattern aligns with observed friendship networks in our data, suggesting that peer effects operate primarily within same-gender interactions, as discussed in Section 2.3.1. These findings highlight the importance of accounting for both peer and gender composition when evaluating the effects of classroom environments on cognitive development.

³Classroom and teacher quality vary but are orthogonal to peer quality, as teachers were also randomly assigned within schools and grades.

While we cannot conclusively distinguish between different mechanisms, our findings suggest that the negative effects of exposure to high achievers are concentrated in environments with heightened competition, greater peer interaction, and increased salience of relative class rank. First, the adverse effects are largest for high-performing children, whose math test scores decrease by 0.030 SD and who experience declines in happiness. These results hold after controlling for classroom rank, suggesting that the effects are not driven by mechanical reductions in students' relative rank within the class. Second, the negative effects are more pronounced in classrooms with more high achievers (on average, eight) and in smaller schools, particularly those with many high achievers.

We also analyze how teacher quality mediates peer effects. Using detailed observational data on teacher behaviors, we find that higher-quality teachers, particularly those skilled in classroom organization, can mitigate the detrimental effects. These results highlight the critical role teachers play in lessening the negative impact of high achievers, especially through effective classroom organization and group management. Moreover, the findings underscore the need to understand how peer effects unfold in classrooms with less effective teachers, and the specific channels through which teacher practices mediate peer effects.

A potential concern in estimating peer effects under random classroom assignment is the weak variation in peer composition as classroom size increases (Angrist, 2014). In our setting, identification relies on residual variation in the proportion of high achievers due to finite-sample imbalances in classroom assignments. To assess this variation, we examine whether the number of high achievers assigned to each classroom follows a hypergeometric distribution with school-specific parameters. Only 8.2% of classrooms have a number of high achievers that lies outside the 95% confidence interval implied by this distribution. Furthermore, the 90th and 10th percentiles of the observed distribution are 7 and 2, compared to expected values of 6.7 and 2.9, respectively. These results suggest that the observed variation in peer composition is consistent with theoretical expectations.

Additional diagnostic analyses, described in Section 2.3.1, further support the validity of our empirical strategy.

This paper contributes to the literature in four ways. First, it adds to the experimental evidence on peer composition in learning environments. While previous studies have studied random peer assignment in middle schools (Busso and Frisanchi, 2021), boarding schools (Zárate, 2023) and college settings (Sacerdote, 2001; Carrell et al., 2009, 2013; Sacerdote, 2014; Feld and Zölitz, 2017, 2022), our study focuses on elementary school, where peer exposure is more intense. We leverage a unique experiment with repeated random assignment and nearly perfect compliance to estimate the causal effects of high-achieving peers and their non-linear impacts.

Second, we extend the literature on the effects of high achievers beyond academic achievement (Balestra et al., 2023), admission scores (Busso and Frisanchi, 2021), and long-term outcomes (Cools et al., 2022; Mouganie and Wang, 2020; Bertoni et al., 2020), by studying effects on executive function and non-cognitive skills. Our rich data on teacher quality allow us to analyze how peer effects interact with other classroom inputs. We also find that effects are stronger in environments with heightened competition, greater peer interaction, and increased salience of relative class rank.⁴ While related to class rank effects studied in Carneiro et al. (2025b), our results are not mechanically driven by reductions in students' relative rank alone; instead, both peer composition and rank independently shape student outcomes.

Third, we contribute to the literature on the timing and persistence of the effects of educational inputs (Jacob et al., 2010; Chetty et al., 2014b; Golsteyn et al., 2021; Carneiro et al., 2025b). A novel aspect of our study is the ability to estimate how peer composition effects vary across grades and evolve over time. We show that these effects are largest in early grades, particularly in 1st grade, and decline significantly as children age.

⁴See de Roux and Riehl (2022); Chen and Hu (2024) for a review of university settings

Finally, our findings inform student-grouping policies, especially in developing countries where peer effects may differ due to larger class sizes, tighter resource constraints, and institutional settings.⁵ We provide new evidence from elementary schools in Ecuador, underscoring the importance of accounting for context-specific peer dynamics and institutional settings when designing grouping policies to improve student outcomes (Mouganie and Wang, 2020; Busso and Frisanchi, 2021; Chen and Hu, 2024).

The remainder of the paper is organized as follows. Information about Ecuador, the experimental setting, and data can be found in section 2.2. Section 2.3 outlines the empirical strategy. Section 2.4 provides the main results, the dynamics of the effects, and heterogeneity analysis. Conclusions and policy discussion can be found in section 2.5.

2.2 Setting and Data

2.2.1 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 ages 5 to 14. The education system is divided into elementary school (Kindergarten through 6th grade), middle school (7th through 9th grades), and high school (10th through 12th grades). The school year follows a dual-calendar system. In the coastal region, it runs from May to February (similar to most countries in the Southern Hemisphere and many in Latin America), while in the highlands and Oriente (eastern) regions, it runs from September to June (similar to most countries in the Northern Hemisphere).⁶

⁵See Sacerdote (2011, 2014) for a review of previous studies.

⁶According to the 2010 census, 53% of the population lived on the coast, 42% lived in the highlands, and 5% in the Oriente region.

Approximately 4.4 million children were enrolled in the education system during the 2012–2013 school year, with 78 percent attending public schools.⁷ Ecuador has made significant progress in expanding access to education. In 2013, the primary school net enrollment rate was 95%, with a completion rate of 97%.⁸ 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 primary school children achieved the minimum proficiency level in mathematics, with a slight gender gap—49.8% for males compared to 46.7% for females.⁹ 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 both the quality of the education system and the instruction provided to students.¹⁰ However, key challenges persist. For example, the pupil-qualified teacher ratio remained stable at around 27.5 during this period.¹¹

2.2.2 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

⁷Data obtained from <https://educacion.gob.ec/datos-abiertos/> on November 4th, 2024

⁸The 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.

⁹This 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 were obtained from the World Development Indicators (<https://datatopics.worldbank.org/world-development-indicators/>) on November 4th, 2024

¹⁰This corresponds to the fraction of teachers with the minimum academic qualifications required for teaching in primary school.

¹¹All data were obtained from the World Development Indicators (<https://datatopics.worldbank.org/world-development-indicators/>)

impacts cognitive and non-cognitive skills.¹² Each school had at least two classrooms per grade, with most having exactly two. In the 2012 school year, an incoming cohort of children was randomly assigned to Kindergarten classrooms within each school. These children were then randomly reassigned to classrooms in subsequent grades through 6th grade in 2018. Compliance with the assignment rules was very high, averaging 98.9% (Carneiro et al., 2025b). As a result, children who remained in the same school throughout elementary school were exposed to seven exogenous, orthogonal shocks to classroom composition.¹³

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).¹⁴ 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.4, we cannot reject the null hypothesis that there is no correlation between child i 's achievement and that of her classroom peers, confirming that the random assignment was successful. Further details on the classroom assignment rules, randomization tests, and compliance with randomization are provided in Appendix B.2.

¹²These schools are a random sample of all public schools with at least two Kindergarten classrooms in the Ecuador's coastal region 2012. See Araujo et al. (2016) for more details on the school selection. Moreover, they show that the characteristics of students and teachers in the sample closely resemble those in a nationally representative sample of schools in Ecuador.

¹³It is worth noting that random reassignment may separate students from their friends. However, out-of-class friendship ties are orthogonal to peer composition given the random assignment.

¹⁴We 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, children were ordered by their last and first names and 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.

2.2.3 Data

Child data

At the beginning of Kindergarten, we collected baseline data on maternal education, household wealth, preschool attendance, and children’s vocabulary skills.¹⁵ 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 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).¹⁶ We aggregated correct math responses for each component separately using Item Response Theory (IRT) and then computed the total math score, with each component receiving equal weight.

Additionally, we collected data on executive function for each grade from Kindergarten to 4th grade. Executive function encompasses a set of self-regulatory skills involving various regions of the brain, particularly the prefrontal cortex. It is generally divided into three domains: working memory and attention, inhibitory control, and cognitive flexibility.¹⁷ These skills are crucial for young children to adapt and learn effectively in

¹⁵To 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.

¹⁶The 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 0, and the right end 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)).

¹⁷Working 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 refers to 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.

school, as they are needed to pay attention, take turns, ask questions, remember steps, and solve math problems, among other classroom tasks. More importantly, these skills have been shown to predict long-term outcomes in adulthood related to labor market success, health and crime (Moffitt et al., 2011). We compute scores for each of the three domains, as well as an overall score.¹⁸

In 1st grade, we collected measures of happiness and effort. For happiness, we asked children whether they were always, sometimes, or never happy at school and in their classroom. Similarly, for effort, we asked whether they always, sometimes, or never made an effort to understand what their teachers were teaching and to learn as much as possible in school. Most children provided consistent responses to both questions. We aggregated the responses 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 depression, we used the Patient-Reported Outcomes Measurement Information System (PROMIS) Depression Scale for children aged 11-17, 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 chose 10 questions from the Dweck "Mindset Quiz"; growth mindset refers to the belief that intelligence is malleable and can be developed with effort (Blackwell et al., 2007). Finally, to measure grit, we adapted four questions from the 8-item Grit Scale for children (Duckworth and Quinn, 2009); grit refers to the ability to persevere at a given task. For each outcome, we aggregated responses using factor

¹⁸Unlike 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.

analysis. Further details on the child assessments and the IRT are provided in Appendix B.3.

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

High Achievers Information

At the end of each grade, we asked teachers to identify five students with the highest learning.¹⁹ Teachers did not have access to any test results from the experiment and had to rely solely on their own knowledge of students' academic performance. Therefore, our analysis captures the behaviors and characteristics teachers associate with high learning. To mitigate concerns about subjectivity in these rankings, we classify a student as a high achiever if at least 50% of the teachers who observed them in previous grades identified them as such.²⁰ This approach implies that the definition of a high achiever may evolve over time as students are evaluated by more teachers. Some students may cease to be identified as high achievers, while others may begin to be recognized. By aggregating multiple, independent evaluations, this procedure reduces measurement error from any single teacher assessment, or from individual teacher behavior toward particular

¹⁹Importantly, we did not ask teachers to identify top and bottom performers separately in math and language.

²⁰For example, a child in 4th grade identified as having the highest learning by three former teachers would be considered a high achiever for that year. More importantly, teachers appear only once in our sample and are not allowed to teach in a future grade.

students, and provides a stronger signal of student performance in later grades.²¹ Using this information, we construct the leave-one-out proportion of high achievers in each classroom.

There are several alternative ways to identify high achievers. For instance, one could use measures of innate ability (e.g., IQ scores), baseline characteristics (e.g., PPVT scores), parental education (Cools et al., 2022), or placement exam results (Busso and Frisanchi, 2021). However, we rely on teachers' rankings to define high achievers for three main reasons. First, this approach enables us to identify high achievers dynamically each year—something that pre-experiment measures do not consistently allow. While baseline scores may provide valuable insights, they are not available for all children in the sample. For example, PPVT scores, though informative, are available for only about 14,000 of the nearly 28,000 students in our sample. Second, because tests in each subsequent grade incorporate both previously asked and new questions, they induce serial correlation in test scores over time, which could mechanically classify students as high achievers across multiple years. Finally, teacher rankings reflect classroom performance rather than performance on standardized tests, allowing them to serve as a distinct measure of "ability". Teachers may also value other characteristics beyond academic performance (e.g., persistence, grit) that would not be captured by test scores or other measures.

Nevertheless, relying on teachers' rankings to define high achievers presents some challenges. First, a key requirement for our empirical strategy is that the definition of high achievers does not suffer from reflection problems (Manski, 1993). In our context, teachers have information about the children when ranking them and can compare students within the same classroom, which could potentially bias their selection of the

²¹Given concerns about potential gender bias, Appendix Table B.1 shows that female and male students identified as high achievers are similar in baseline characteristics. This suggests that teachers apply similar criteria when identifying high-achieving students of different genders. Furthermore, our main results hold when using alternative definitions of high achievers based on test scores as discussed in Section 2.4.2.

five highest-performing students. To mitigate this concern, we rely only on information from teachers in previous grades and do not use any data from the current grade that could be affected by reflection issues. Second, the probability of being designated as a high achiever is higher in smaller classrooms. To address this, we aggregate information from multiple teachers across different grades, which helps smooth out any idiosyncratic variation due to class size. Moreover, our empirical strategy compares students across classrooms within the same school-grade. Due to the experimental design, classroom sizes within a school-grade are very similar, so the probability of being identified as a high achiever should also be comparable across classrooms.

Teacher and Classroom data

We use the Classroom Assessment Scoring System (CLASS) (Pianta et al., 2015) to measure teacher behaviors based on recordings of classroom interactions over the course of an entire school day. The CLASS measures teacher behavior across three domains: Emotional Support, Classroom Organization, and Instructional Support.²² Within each domain, there are three to four CLASS dimensions, each rated on a scale of 1 to 7. The CLASS protocol provides coders with specific guidance for assigning scores, categorizing them as "low" (1–2), "medium" (3–5), or "high" (6–7). The behaviors coders assess within each dimension are clearly defined and highly 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

²²Emotional Support measures how teachers promote a positive classroom climate and respond to students' emotional needs. Classroom Organization refers to the routines and procedures teachers use to manage students' behavior, time, and attention. Instructional Support evaluates how effectively teachers promote students' thinking through their implementation of the curriculum.

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 informed of the filming only on the day itself, ensuring that the videos captured typical, planned interactions with students. Each classroom has only one teacher that teaches the same class during the calendar year, without a classroom aide, who was responsible for all academic subjects (excluding physical education, 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.²³ Further details on the CLASS dimensions, the scoring protocol, and its application in Ecuador are provided in Appendix B.4.

One disadvantage of the CLASS is that it is not a cardinal scale, which means it cannot be used directly 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). Following Araujo et al. (2016), we use the CLASS as a measure of Responsive Teaching. We categorize teachers into two groups (low and high quality) based on their CLASS scores, separately for each domain and for the total score.

2.2.4 Descriptive Statistics and Preliminary Checks

Table 2.1, Panel A provides summary statistics for the children in the sample. On average, children were five years old on the first day of Kindergarten, with 60% having attended preschool, and half being girls. The average score on the vocabulary skills (PPVT) test is 83 points, with a considerable range from 55 to 145. Furthermore, 90% of

²³In cases where the two coders provided significantly different scores, a third coder was assigned to evaluate the teacher. The inter-coder reliability was approximately 0.92, suggesting that measurement error due to from substantial differences across coders was relatively low.

the children self-report being happy, and 85% report putting in effort. Thirteen percent of children are classified as high achievers by their previous teachers. Table 1.1, Panel B presents information on household and parental characteristics. The average age of mothers is 30 years and 34 years for fathers, with both having completed approximately 8.5 years of education. Regarding household conditions, 83% report having piped water, and 43% have a toilet at home.

Table 2.1, Panels C and D present classroom and teacher characteristics, respectively. On average, 15% of a student's peers in a given classroom are classified as high achievers based on teacher rankings, and the average classroom size is 37 students. However, these proportions vary significantly across classrooms, ranging 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 each assessment domain.

Table 2.1: 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.

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 B.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. In fact, more than 60% of the classrooms had three or more high achievers. As a result, only 26 classrooms did not contain a high achiever.²⁴

Second, one potential concern with our preferred measure of high achievers is whether teachers can effectively identify them. Table 2.2 provides summary statistics of the characteristics of the children in our sample, comparing those classified as high achievers, based on the teacher's rankings from previous grades, with those who are not. The table shows that children classified as high achievers have higher math and executive function scores, as well as higher baseline vocabulary test (PPVT) scores. They are more likely to be female and report higher levels of happiness and effort. Overall, there are notable differences in socioeconomic status between high achievers and other children; for instance, high achievers tend to 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 high achievers correctly, at least based on their observable characteristics, while also considering other relevant characteristics.

²⁴This 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.

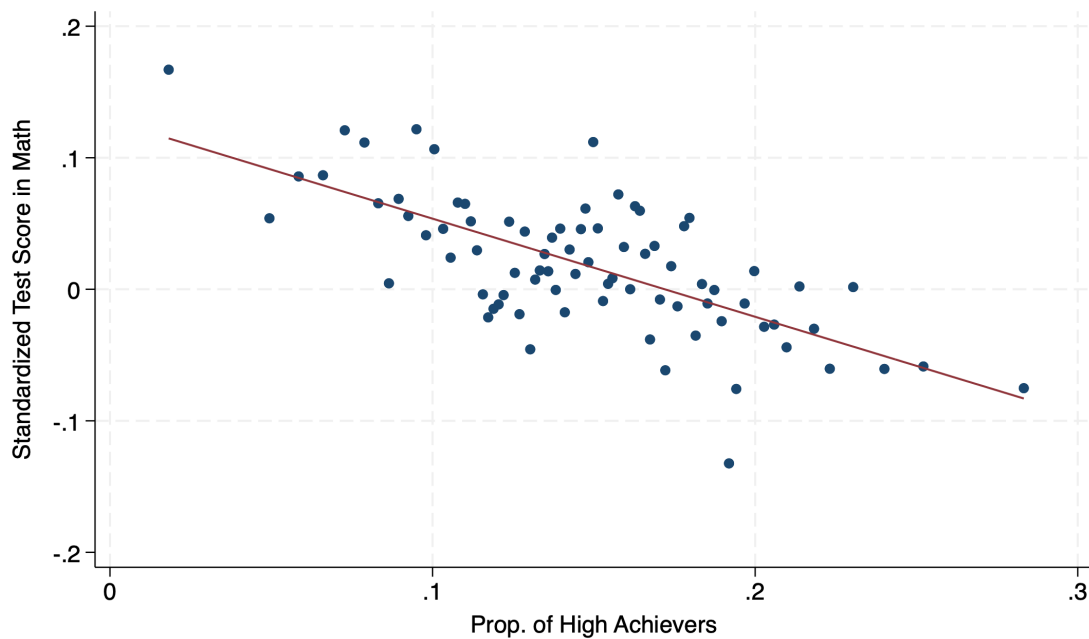
Table 2.2: Characteristics of high-achieving students

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

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

Finally, Figure 2.1 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 random assignment, without considering other factors like ability or gender. The figure suggests that children exposed to more high achievers obtained lower math test scores.

Figure 2.1: 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.

2.3 Empirical Strategy

2.3.1 Main Model

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

$$y_{icst} = \beta_1 \text{FracHighAchievers}_{cst} + \mathbf{X}_{icst}\alpha + \delta_{st} + \varepsilon_{icst} \quad (4)$$

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 . $\text{FracHighAchievers}_{cst}$ 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 respond to high achiever peers depends on the gender of those high achievers (Busso and Frisanchi, 2021; Cools et al., 2022). We explore this question by analyzing how test scores differ for girls and boys when they are exposed to high-achieving peers, separated by gender. Therefore, we estimate the following model:

$$y_{icst} = \beta_1 \text{MaleFracAchievers}_{cst} + \beta_2 \text{FemaleFracAchievers}_{cst} + \mathbf{X}_{icst}\alpha + \delta_{st} + \varepsilon_{icst} \quad (5)$$

where $\text{MaleFracAchievers}_{cst}$ ($\text{FemaleFracAchievers}_{cst}$) 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 classrooms within the same school and grade, which is a common approach in the literature. The key identification assumption is that, conditional on school-by-grade fixed effects, classrooms with differing peer composition do not differ systematically along other dimensions. The school-by-grade fixed effects account for differences across cohorts within a given school and for school-grade characteristics that are constant across students. This key assumption is supported by the random assignment of students to classrooms at the beginning of each grade.

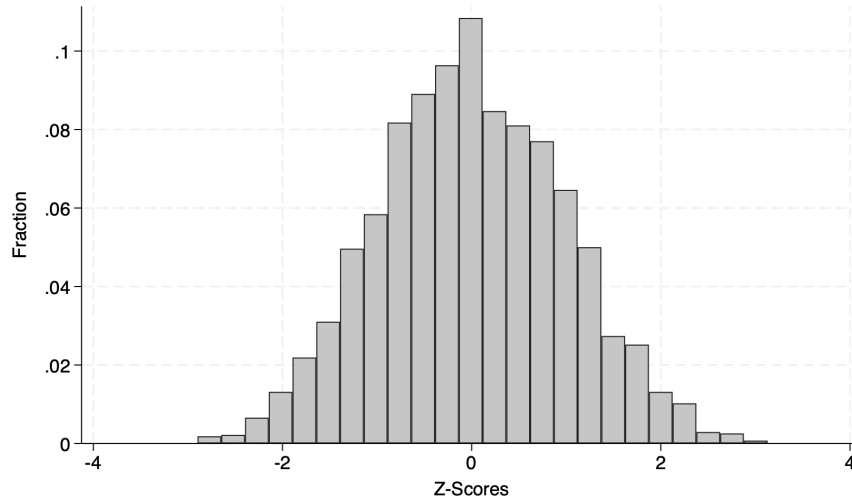
A potential concern is weak variation in peer composition, which can arise when

estimating peer effects under random assignment to classrooms (Angrist, 2014). This issue is particularly relevant in settings with large peer groups, where randomization may produce limited variation. In our context, the variation in the proportion of high achievers decreases with class size, and identification relies on residual variation due to finite-sample imbalances in classroom assignments. However, this concern is mitigated in our setting, as most classrooms have fewer than 45, with an average of 37.

However, recognizing these potential concerns, we conduct four diagnostic analyses to assess the extent of variation in the proportion of high achievers and to test for random assignment. First, we assess whether the number of high achievers per classroom follows a hypergeometric distribution with school-specific parameters. Figure 2.2 shows the distribution of the standardized deviations between the observed and expected number of high achievers under this distribution. Only 8.2% of classrooms have a number of high achievers that lies outside the 95% confidence interval implied by the hypergeometric distribution. Moreover, the 90th and 10th percentiles of the observed distribution are 7 and 2, compared to expected values of 6.7 and 2.9, respectively. These results suggest that the observed variation in peer composition is consistent with theoretical expectations under random assignment.

Second, the ability to exploit peer composition relies on sufficient residual variation in the key variables after accounting for fixed effects. Given the random assignment, the distribution of the proportion of high achievers in a given classroom, conditional on school-by-grade fixed effects, should be approximately normal. Appendix Figure B.2 shows that the distribution of the residualized proportion of high achievers across classrooms is indeed approximately normal, consistent with random assignment. This suggests that variation in peer composition is likely arbitrary. More importantly, the figure also shows some within school-grade variation in the proportion of high achievers, further supporting the validity of our empirical strategy.

Figure 2.2: Distribution of Standardized Deviations from Hypergeometric Expected Number of High Achievers



Notes: The figure above shows the distribution of the standardized deviations of the observed number of high achievers in the classroom from its expected value under the hypergeometric distribution with school-specific parameters.

Third, to obtain precise estimates, there must be sufficient variation in the leave-one-out proportion of high achievers after controlling for school-by-grade fixed effects. Appendix Table B.2 reports the variation in this proportion before and after including the fixed effects. As expected, including the fixed effects reduces the standard deviation by approximately 30 percent. Nevertheless, the remaining variation is substantial, suggesting that there is sufficient identifying variation to credibly estimate the effects.

Finally, we test whether the variation in the proportion of high achievers within a school-grade is consistent with random assignment by comparing the actual distribution to a simulated one. To this end, we conduct Monte Carlo simulations in which students are randomly assigned to classrooms within the same school and grade. Following Bietenbeck (2020), we take the number and size of classrooms, as well as the number of high achievers

per classroom, using the observed data. In the simulated dataset, we regress the proportion of high achievers on school-by-grade fixed effects and collect the residuals. Appendix Figure B.3 plots the distribution of these residuals from 1,000 replications alongside those from the observed data. The two distributions are visually similar, which is consistent with random assignment and further supports the assumption that high achievers were randomly assigned 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 alternative specifications: i) we exclude the grade fixed effect; ii) we include separate school and grade fixed effects; and (iii) we include teacher controls to account for the possibility that teachers may influence the classroom environment and affect student outcomes. As we show later, the results from these specifications are similar to those obtained using the main specification.

Second, we test the sensitivity of the estimates to different ways of defining who is classified as a high achiever in the classroom. For our preferred measure, we estimate three new proportions by classifying a child as a high achiever 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 achievers using only the information from the previous year's teacher, instead of using information from all teachers. Given the availability of test scores, we also construct two measures based on the children's performance: (i) whether a student is above the 95th percentile of the corresponding test score from the previous year, and (ii) whether a student is above the 95th percentile of the baseline PPVT test. The results show no significant changes.²⁵

²⁵We also use the parents' education to construct a proportion of high achievers, similar to the approach

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 B.4 shows 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 centers 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. A potential concern is that differential attrition could present an estimation challenge. Appendix Table B.11 shows the impact of the leave-one-out proportion of high achievers on the likelihood of leaving the sample between two consecutive grades. The table shows that children are no more likely to attrit when exposed to a higher proportion of high achievers. Therefore, we find no evidence of selective attrition. Nevertheless, in Appendix Table B.12, we restrict our sample to the balanced panel of children and estimate the main equation. We find that the results are similar.

2.3.2 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:

$$y_{isc,t+l} = \beta_1 \text{FracHighAchievers}_{sct} + \mathbf{X}_{isct} \alpha + \delta_{st} + \varepsilon_{isct} \quad (6)$$

used by Cools et al. (2022). Appendix B.5 shows that the results are consistent with those obtained using the preferred definition.

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 (1st grade to 6th grade). When $l = 0$, equation 6 is equivalent to equation 4 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 6 provides estimates of the medium-term effects of peer composition at various lags (at most 6 for Kindergarten), which we label $\beta_{t,l}$.

2.4 Results

2.4.1 Main Results

We estimate the regression model specified in Equation 4 separately for each cognitive and non-cognitive test score. Table 2.3 shows that a higher proportion of high-achieving peers reduces scores in math and executive function tests. For instance, 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 of one standard deviation is associated with a 0.014 standard deviation decrease in executive function scores.²⁶ These results are surprising given the previous literature on high-achieving peers, which generally finds positive effects. However, they are consistent with the findings of Hoxby and Weingarth (2005). In terms of magnitude, the effects are slightly smaller than the 0.04SD decrease in GPA found by Chen and Hu (2024). Their estimates, however, focus on high-ability students, which may explain the difference, as we discuss later.

²⁶Appendix Table B.3 shows the results for each of the components included in the math and executive function tests. It indicates that the proportion of high achievers negatively affects all components of math, with stronger effects on number sense and word problems. In the executive function components, only cognitive flexibility and working memory are negatively affected, while inhibitory control is not.

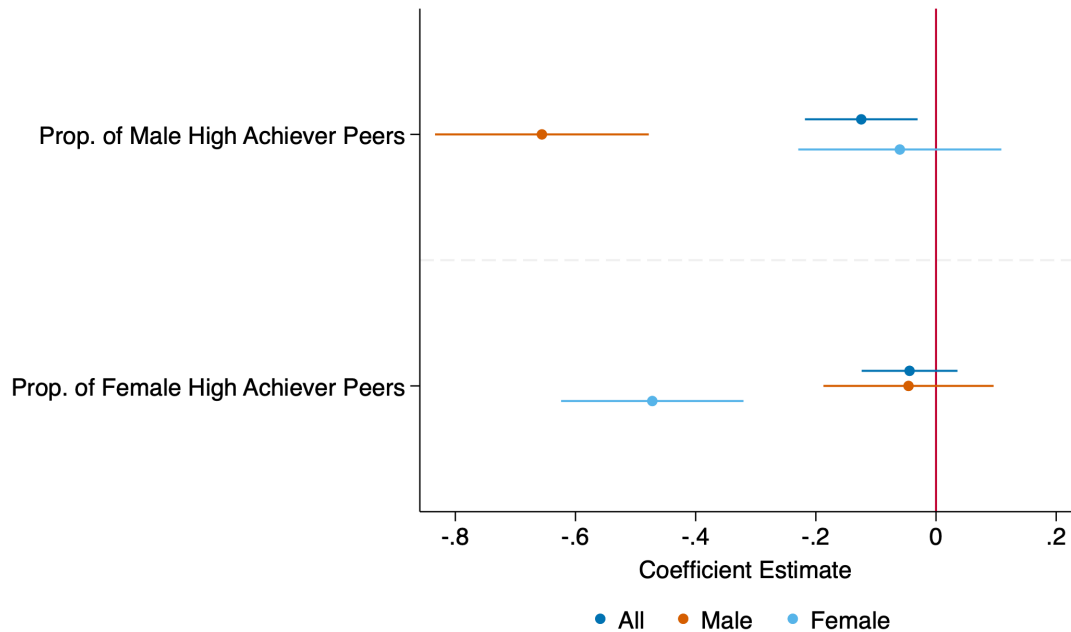
Table 2.3: Effects of High Achievers on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of High Achiever Peers	-0.156** (0.065)	-0.194** (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.

There are three plausible explanations for our findings. First, the presence of high-achieving peers may lower the relative academic rank of other students within the classroom. Students who are pushed down to a lower rank in the local distribution may reduce their effort or lose self-confidence, which can negatively affect their performance, as documented in tracking settings (de Roux and Riehl, 2022). Second, the presence of high achievers may intensify competition, which could negatively impact students’ mental health, motivation, and stress levels, in line with previous evidence on the effects of competition in education (Chen and Hu, 2024). Finally, teachers may focus more attention on high achievers and neglect others students, or they may adjust their teaching strategies and classroom management, potentially mediating the effects on learning outcomes.

Figure 2.3: Heterogeneous Effects of High Achievers on Math Scores 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.

Given the asymmetric gender effects of peers found in the literature, Figure 2.3 shows that, on average, the coefficients for the proportion of male high achievers tend to be larger in magnitude than those for females, although the differences are not statistically significant. The figure also illustrates how the share of female and male high achievers affects male and female students separately. Interestingly, the results suggest that the proportion of male high achievers affects only male students, while the proportion of female high achievers only female students. These findings are consistent with some of the results in Lavy et al. (2012), who found that high-achieving male peers negatively impact boys, and they suggest that crowding out in top-tier activities may drive these effects.

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

One potential concern is mean reversion, where children who initially score lower might experience larger increases in the next grade, while those with higher initial scores could experience declines. However, given that teachers identify high achievers without knowing their test scores, mean reversion is unlikely to explain our results for two reasons. First, our preferred measure of high achievers relies on rankings provided by multiple teachers, not just one, so no explicit design allows teachers to strategically adjust who is considered a high achiever. Second, as explained before, teachers, on average, do not make mistakes when identifying high achievers. Moreover, Appendix Figure B.6 shows no evidence suggesting that teachers across grades differ their ability to identify higher-achieving students. The distribution of test scores for high achievers is consistently to the right of that for non-high achievers.

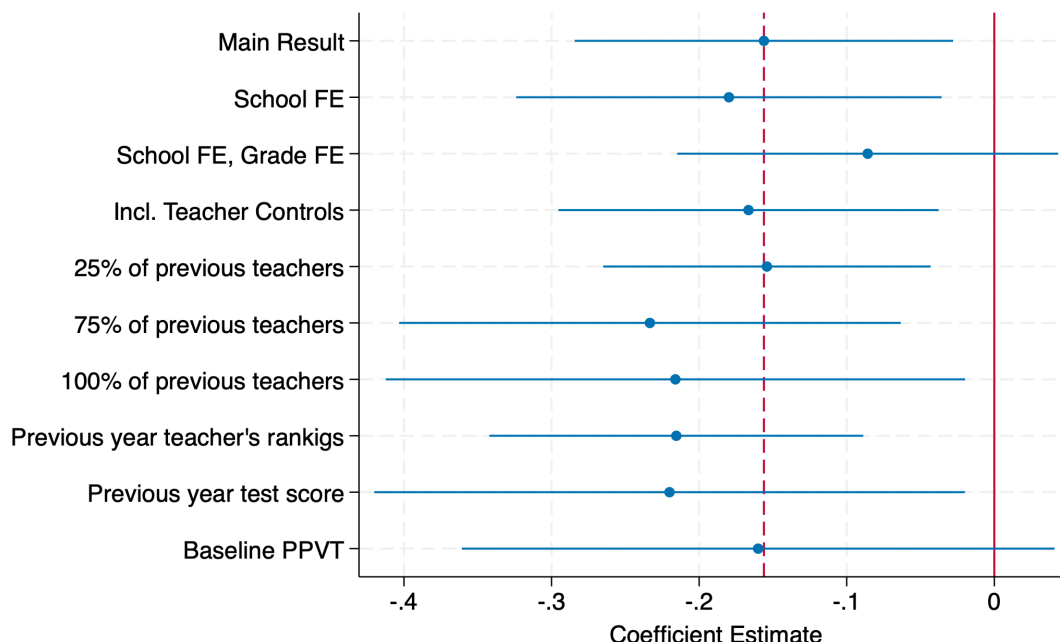
However, recognizing that there could be a mechanical reason for why test scores might decrease, we address mean reversion by analyzing changes in the distribution of children at the top end. Over the sample period, math test scores exhibit some mean reversion: children who were in the top 20% of scores 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, 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.

2.4.2 Robustness checks and Alternative Specifications

This section provides a brief summary of several alternative specifications and robustness checks to address concerns about the main specification. The first 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 school and grade fixed effects separately. Finally, the last specification check includes teacher controls to account for how teachers may influence the classroom environment and affect student outcomes through their interactions. Figure 2.4 presents 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.

Figure 2.4: Robustness Checks on Math Scores



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 1.5. 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.

The next robustness checks test the sensitivity of the estimates to different ways of defining who is classified as a high achiever in the classroom. First, our preferred measure uses information from all the teachers in previous grades and classifies a child as a high

achiever 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 the results are to the use of other cutoffs. Second, we construct a proportion of high achievers using information from only the previous grade teacher, rather than from all the previous teachers. The next four estimates in Figure 2.4 show the main specification using cutoffs of 25%, 75%, and 100%, 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 achievers using test scores. We construct two different measures: (i) whether a student is above the 95th percentile of the corresponding test score from the previous year, and (ii) whether a student is above the 95th percentile of the baseline PPVT test. The last two coefficients in Figure 2.4 show the results for these measures. The coefficient using the previous year's test score shows similar results, while the one using the PPVT score is similar in magnitude but not statistically significant. However, this coefficient should be interpreted cautiously, as the PPVT was collected only at the baseline, and new entrants do not have information on this test. Additionally, since the PPVT is a vocabulary test rather than a math test, this could explain the absence of effects on math scores.

2.4.3 Dynamics

We estimate the dynamics of the effects by estimating the regression of peer composition separately for children in each grade (1st grade to 6th grade), both contemporaneously (without lags, as in equation 4 above) and at various lags (up to 5, as in equation 6). Table 2.4 presents the results for the proportion of high achievers in math. Column (1) shows that the short-term effects of the proportion of high achievers on math scores are negative only for 1st grade.

Table 2.4: Cumulative Effects of High Achievers on Math Scores

	Lags					
	0	1	2	3	4	5
	(1)	(2)	(3)	(4)	(5)	(6)
1st grade	-0.521** (0.205)	-0.172 (0.120)	-0.227** (0.112)	-0.128 (0.107)	-0.155* (0.0911)	-0.150* (0.0851)
2nd grade	-0.077 (0.160)	-0.121 (0.123)	-0.133 (0.0870)	-0.130* (0.0713)	-0.066 (0.0772)	
3rd grade	-0.172 (0.158)	-0.317*** (0.115)	-0.082 (0.0814)	-0.107 (0.0846)		
4th grade	-0.089 (0.154)	-0.140 (0.0972)	-0.124 (0.0850)			
5th grade	-0.079 (0.125)	0.019 (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.

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 coefficient in column (2) represents the impact of peer composition in 1st grade on second-year scores, while column (3) represents the impact of peer composition in 1st grade on third-year scores. These coefficients are relatively

stable and statistically significant. However, in contrast to previous evidence on ranking (Carneiro et al., 2025b), the effects over time decline. Interestingly, the table suggests that only 1st grade matters, and peer composition during this early grade appears particularly relevant. These results align with previous evidence on how Kindergarten experiences have long-term consequences (Rury, 2025).

Table 2.5: Cumulative Effects of High Achievers on Executive Function

	Lags			
	0	1	2	3
	(1)	(2)	(3)	(4)
1st grade	-0.270* (0.160)	-0.565*** (0.176)	-0.167 (0.154)	-0.513*** (0.152)
2nd grade	-0.249* (0.146)	-0.170 (0.173)	-0.289** (0.121)	
3rd grade	0.160 (0.168)	-0.529*** (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.

Similarly, Table 2.5 presents the results on executive function, showing similar findings to math test scores for 1st grade, but also indicating effects in subsequent grades. These results suggest that exposure to high achievers has long-term effects on executive function, a set of skills related to a child's ability to plan, focus attention and remember instructions.

2.4.4 Potential Mechanisms

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

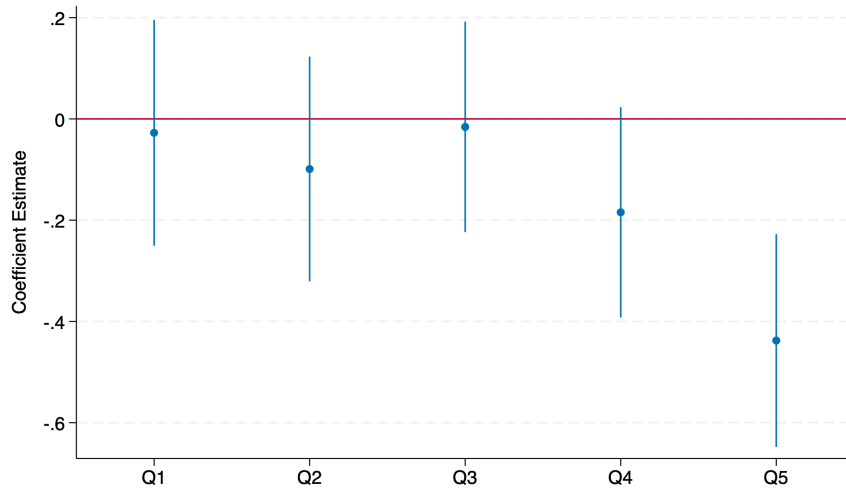
Classroom Environment

In this section, we explore the roles of competition, relative ranking, and classroom environment behind the observed negative peer effects. First, we examine heterogeneity by previous achievement to test whether the effects are concentrated among students at the top of the distribution. Second, we assess whether these effects are driven by changes in classroom rank. Third, we estimate non-linear effects in exposure to high achievers, which may imply increasing competitive pressure. Fourth, we analyze heterogeneity by school size as a proxy for differences in classroom environment. Overall, we find that negative peer effects are more pronounced in environments with heightened competition, increased peer interaction, and stronger awareness of relative ranking, though we cannot conclusively distinguish between these mechanisms.

In our context, children with higher prior-year scores are more likely to experience downward shifts in relative rank. Figure 2.5 divides the sample into quintiles based on

prior math scores and shows that negative peer effects are concentrated in the top quintile. This is consistent with the “Invidious Comparison Model”, which predicts stronger effects among students near the top of the distribution. These findings also align with prior literature showing that high-achieving students tend to be more competitive than their lower-achieving peers (Syeda and Khalid, 2012), and that high-ability students are particularly negatively affected by the presence of other high achievers due to intensified competition among them (Chen and Hu, 2024).

Figure 2.5: Heterogeneous Effects of High Achievers on Math Scores by Previous Test Results



Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers on math test scores by quintiles of the test result in the previous year. 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.

The presence of high achievers could push some students down in the classroom rank distribution, which may reduce effort or self-confidence. Since classroom rank has positive effects on student outcomes in this setting (Carneiro et al., 2025b), we estimate models

that include both classroom rank and the proportion of high achievers. Table 2.6 shows that, for the full sample, the effect of peer composition is statistically not significant once rank is accounted for. However, among students in the top quintile of prior performance, the proportion of high achievers still has a negative effect even after controlling for rank. In particular, Column 4 shows negative effects even when comparing two children with the same classroom rank but exposed to different proportions of high achievers. Appendix Figure B.5, which includes rank in a non-linear way, shows that both peer composition and rank independently shape student outcomes, particularly among top performing students. While Carneiro et al. (2025b) highlight the importance of rank effects, our findings suggest that peer composition also affects test scores.

Table 2.6: Effects of High Achievers and Classroom Rank on Math Scores

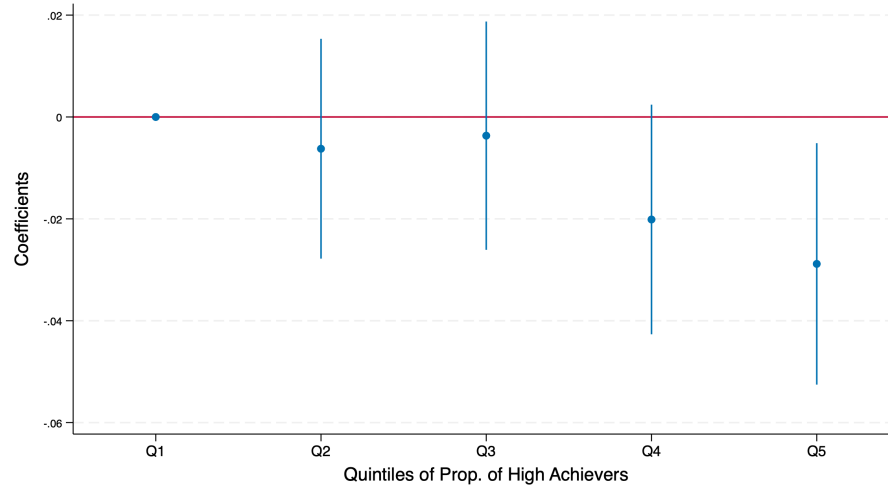
	Full Sample		Top Performing (Q5)	
	(1)	(2)	(3)	(4)
Prop. of High Achiever Peers	-0.071 (0.066)	-0.070 (0.066)	-0.409*** (0.110)	-0.419*** (0.108)
Students' Classroom Ranking	0.241*** (0.028)		0.094 (0.075)	
Mean of Dependent Variable	0.0193	0.0192	1.133	1.134
Observations	87303	87251	17880	17832
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes
Classroom Rank FE	No	Yes	No	Yes

Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers and the classroom rank on math test scores. For Columns 1 and 3, classroom rank is defined as a measure in percentiles (so rank is normalized to be between zero and one) similarly to Carneiro et al. (2025b). For Columns 2 and 4, we include individual fixed effects for each percentile of the classroom rank. 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.

To estimate the non-linear effects of peer composition, we create quintiles based on the proportion of high achievers in the classroom. Figure 2.6 shows that negative effects on test scores emerge only in the upper two quintiles. Being in a classroom with fewer than three high achievers has little impact, but being in a classroom with more appears detrimental. This pattern suggests that competition may play a role, as students might feel increased pressure to perform or experience crowding-out effects due to limited opportunities for top-tier activities or recognition. Overall, these results indicate that

the additional competition produced by an increase in the proportion of high achievers negatively impacts test scores.

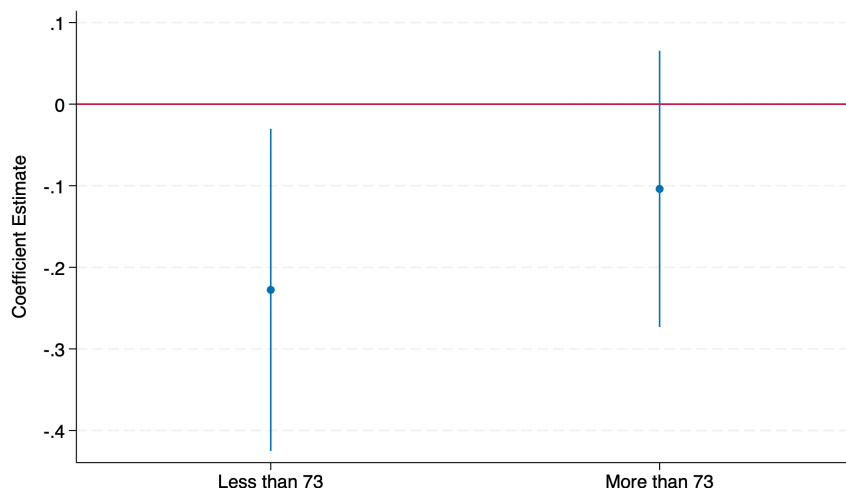
Figure 2.6: Non-Linear Effects of High Achievers on Math Scores



Notes: The figures above report non-linear estimates from regressions of the quintiles 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. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Furthermore, we estimate effects by school size, using the average number of students per school over the seven years of the experiment. We use the median to define small and large schools. We expect peer effects to be stronger in smaller schools as they are characterized by heightened competition, greater peer interaction, and increased salience of relative class rank. Figure 2.7 shows that the detrimental effects are concentrated in smaller schools, with estimates for small and large schools statistically different. However, as discussed earlier, variation in peer composition decreases with class size. Therefore, the greater variation observed in smaller schools could increase the likelihood of detecting effects.

Figure 2.7: Heterogeneous Effects of High Achievers on Math Scores 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 math test scores. Schools are separated into two groups: small and large, using the median size as the cutoff which corresponds to 73 students. 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.

Finally, we analyze the interaction between peer composition and school size. Appendix Figure B.7 shows that the non-linear effects are more pronounced in small schools (Panel A), but are not detectable in larger schools (Panel B). In particular, students in small schools exposed to the highest proportions of high achievers experience the largest declines in test scores. These findings reinforce the idea that the classroom environment mediates the effects, though we cannot conclusively identify the underlying mechanism.

Effects on other outcomes

In the previous section, we show that peer effects are mediated by classroom environment. In this section, we explore how high achievers affect other outcomes, particularly mental health, non-cognitive skills and motivation. First, Table 2.7 shows no significant

effects on mental health outcomes and non-cognitive skills, such as depression, self-esteem, grit, and growth mindset in 6th grade. However, these estimates should be interpreted cautiously, as the smaller sample size leads to noisier estimates, and our power to detect effects is reduced.

Table 2.7: 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 (0.258)	-0.057 (0.254)	-0.049 (0.262)	-0.136 (0.236)
Mean of Dependent Variable	0.001	0.001	0.001	0.001
Treatment Effect of 1SD	-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.

Now, we turn to self-reported measures of happiness and effort. One potential explanation for the test scores results is that children feel less happy when they are exposed to more high achievers, leading them to put in less effort on their work and tests. If students exert less effort, the estimates of the effects on test scores would need to account for that. Table 2.8 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. Interestingly, Appendix Figure B.8 further shows that the decline in happiness is concentrated among students who were at

the top of the distribution in the previous year. These results suggest that reductions in test scores could be attributed to decreased motivation. While we do not find evidence of reduced self-reported effort, the drop in happiness may reflect a decline in motivation or increased stress, which could affect their performance. Together with the evidence on rank and peer composition, these results suggest that negative peer effects may partly operate through reductions in well-being and motivation, especially among top performing students.

Table 2.8: Effects of High Achievers on Happiness and Effort

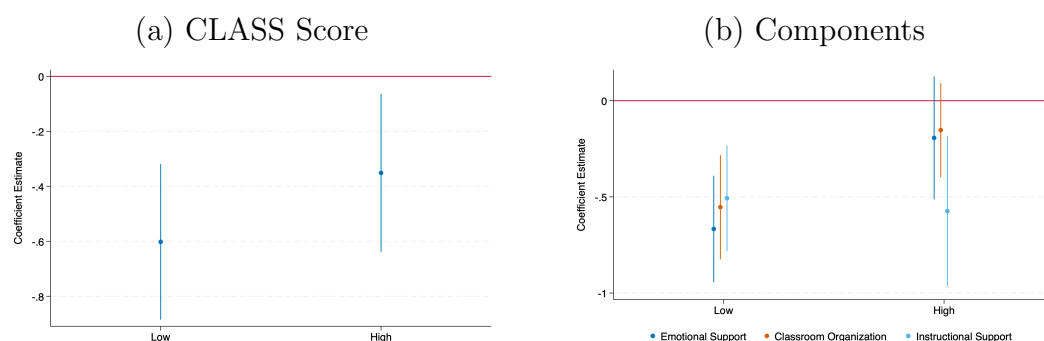
	Happiness	Effort
	(1)	(2)
Prop. of High Achiever Peers	-0.098** (0.048)	0.019 (0.058)
Mean of Dependent Variable	0.903	0.854
Treatment Effect of 1SD increase	-0.008	0.002
Observations	12034	12034
Controls	Yes	Yes
School-by-grade FE	Yes	Yes

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.

Teacher Quality

Given the existing literature on teacher quality and its positive effect on test scores, we test whether teacher behaviors, as measured by the CLASS, can mitigate the negative peer effects. Panel A in Figure 2.8 shows that high-quality (those above the median) teachers can reduce the negative effects, although some detrimental consequences on math test scores remain. Panel B in the same figure examines the results by CLASS domain and illustrates that teachers who are high-quality in emotional support and classroom organization are the ones most capable of mitigating these effects. These findings suggest that teachers with better-planned classroom routines and structure may be better equipped to distribute attention more equitably and manage peer dynamics effectively. Overall, these results highlight the importance of teacher quality, highlighting the role of classroom management in shaping peer effects on test scores.

Figure 2.8: Heterogeneous Effects of High Achievers on Math Scores by Teacher Quality



Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers on math test scores by teacher quality measured using the CLASS (*Classroom Assessment Scoring System*) score. Teachers are separated into two groups: low and high quality, using the median score as the cutoff. 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.

2.5 Conclusion

This paper analyzes the impact of peer and gender composition on students' cognitive and non-cognitive skills in elementary school settings. The data come from a unique longitudinal experiment in Ecuador where children are randomly assigned to classrooms at the beginning of each school year for seven consecutive years. This random assignment implies that the proportion of high achievers in a classroom within a school is random. Compliance with the random assignment was nearly perfect, averaging 98.9% over the seven years. Notably, in our data, two students with the same underlying ability who attend the same school can experience different peer compositions because they are randomly assigned to different classrooms, each 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. Moreover, these negative effects are stronger 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 and smaller classrooms, suggesting that environments characterized by heightened competition, increased peer interaction, and stronger awareness of relative ranking 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 impact non-cognitive outcomes. These findings contribute to the understanding of 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 leverage the random assignment in elementary schools to 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 present in other peer effect studies. However, we also provide evidence addressing the concerns of weak variation when estimating peer effects under random assignment to classrooms. Second, we show that peer composition has long-term effects, though these effects decline over time. Our results highlight the importance of 1st grade peers, suggesting that peer composition in early grades is particularly relevant. Third, we explore both classroom environment and teacher quality as potential mechanisms behind the observed effects of exposure to high-achieving peers, providing insights into the complex dynamics in early educational settings.

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

Chapter 3

The Impacts of Terrorist Attacks on Socioemotional Skills and Risky Health Behaviors

3.1 Introduction

Exposure to armed conflict events is common for individuals in developing and some developed countries. Over 43 countries around the globe have experienced civil war and violence since 2000 (Marshall, 2020).¹ Events such as terrorist attacks are a canonical example of an Adverse Childhood Experience (ACE) due to their potential impact on health and well-being in the short and long term.² The consequences are likely even more dramatic when exposure occurs during childhood, which is a critical period for developing

¹Moreover, real-time data on the location of armed conflict events around the world suggests that these events are still prevalent, with over 10,000 events and 19,571 fatalities in July 2023 (Raleigh et al., 2010), and that global peacefulness has deteriorated in the last years (Institute for Economics and Peace, 2023).

²See Abadie and Gardeazabal (2003) and Brodeur and Yousaf (2022) for the economic consequences of conflict

essential cognitive and non-cognitive skills (Almond et al., 2018; Conti and Heckman, 2012; Cunha and Heckman, 2007).³ Similarly, armed conflict has been shown to have negative consequences on the accumulation of noncognitive skills in adolescence and worse health outcomes in adulthood (Akbulut-Yuksel, 2014; Jürges et al., 2022).

Traumatic events can affect socioemotional skills and risky behaviors during adolescence as the brain is still developing, and can promote risk-taking and novelty seeking (Giedd, 2004; Winters and Arria, 2011). In fact, socioemotional skills and risky health behaviors are strongly related to future outcomes. For example, socioemotional skills are positively correlated with the labor market and educational results (Borghans et al., 2008; Deming, 2017).⁴ Meanwhile, risky behaviors can have significant consequences on academic performance (Renna, 2008), are often persistent (Chassin et al., 1996; Gruber and Zinman, 2001), trigger other risky behaviors or use of harder drugs (Rees et al., 2001; Grossman and Markowitz, 2005; Deza, 2015), and play a role in premature deaths (Cawley and Ruhm, 2011). Moreover, substance use during adolescence impacts the long-term labor market outcomes and educational attainment (Mezza and Buchinsky, 2021)

In this paper, I study the effect of exposure to terrorist attacks as a child in their municipality on subsequent socioemotional skills and risky health behaviors (drinking and smoking).⁵ To this end, I exploit the within-municipality cross-cohort exposure to terrorist attacks between 2000 and 2012 to identify their causal effects on children's future outcomes in Colombia. The Colombian context offers a valuable opportunity to study the impact of terrorist attacks on teenagers' outcomes because it has experienced a long-lasting conflict spanning six decades, resulting in over 2 million children and

³Early childhood corresponds to the first five years of a child's life and the in-utero period.

⁴See, also, Heckman and Rubinstein (2001); Heckman et al. (2006); Lindqvist and Vestman (2011)

⁵Terrorist attacks include assaults or ambushes against military forces and attacks on civilian buildings and general infrastructure. Armed groups have widely used these attacks to maintain territorial control by generating fear.

adolescents becoming victims of violent events.⁶ The high frequency of such events makes it one of the country's most prevalent forms of adverse childhood experiences. Despite the number of children exposed to violent events, little is known about the medium-term (i.e., adolescence) impacts on their socioemotional skills and risky health behaviors, and most of the evidence is based on descriptive data.⁷

Early-childhood environments affect long-term results (Currie, 2001; Gould et al., 2011; Heckman, 2000; Heckman et al., 2013). However, how exposure to terrorist attacks affects medium-term outcomes is theoretically unclear. On the one hand, exposure to local violence has negative impacts (Koppensteiner and Menezes, 2021). On the other hand, children are resilient and can effectively “bounce back” after traumatic experiences, indicating that exposure to violence may not necessarily have long-lasting effects (Garrett et al., 2019). As a result, the net medium-term effect of terrorist attacks remains ambiguous.⁸ Identifying how the middle years are affected is critical to understanding how the developmental trajectories are shaped by the experiences over the life course (Almond et al., 2018)

This study combines two data sets. The first data set contains individual-level data from the Colombian Longitudinal Survey (ELCA) for 870 teenagers ages 13 to 16 years living in urban areas of the country. This survey includes detailed information on socioemotional skills (Strengths and Difficulties Questionnaire) and self-reported measures of risky health behaviors. The second data set contains information on all terrorist attacks in each municipality in Colombia between 2000-2016. The empirical strategy leverages the within-municipality cross-cohort variation of terrorist attacks. It exploits the

⁶Data available from the National Information Network of the Unit for the Victims Assistance and Reparation. Data retrieved from National Information Network on March 9, 2022.

⁷There is limited evidence on the impact of ACEs and risky health behaviors for other countries (Dube et al., 2006; Shonkoff et al., 2012).

⁸The literature on early childhood interventions shows a similar pattern. Effects on test scores “fade out” by age 8 or 9 but then reappear in adulthood Almond and Currie (2011)

variation in exposure to violence during childhood (0-9 years old) across different cohorts of adolescents within a given municipality. Identifying the effect of municipality violence on future outcomes is challenging, mainly due to sorting concerns. The identification strategy relies on the premise that shocks in the municipality are unpredictable, meaning parents make decisions based on the overall characteristics of the municipality rather than the specific characteristics of their children's cohort.

In the main specification, I include municipality and cohort fixed effects. This set of fixed effects allows us to compare adolescents living in the same area with differential exposure to municipality terrorist attacks during childhood while accounting for national shocks that occurred in a specific year. I find that terrorist attacks deplete the socioemotional skills and adversely impact risky health decision outcomes in adolescents exposed during childhood. Exposure to one more terrorist attack leads to a 0.046 SD increase in the total socioemotional problems, a 0.056 SD increase in the hyperactivity problems score, and a 4.2 percentage points (6.5%) increase in the probability of drinking alcohol with no significant reductions in the age at which they drink for the first time. These results are robust across various specifications, estimation methods, and sensitivity tests (detailed in section 3.5.4).

The identification strategy relies on the assumption that parents only sort across municipalities based on the general characteristics of the municipality, not on the specific characteristics of their child's cohort within the municipality. Several diagnostic analyses support our use of within-municipality cross-cohort variation in exposure to terrorist attacks. First, one potential challenge of the specification above is that the timing of birth could be endogenous. For example, parents who live in more violent municipalities might delay their pregnancy, or they might be able to predict when the municipality is likely to experience more terrorist attacks, which would then affect their fertility decisions. To check whether this is a concern, I provide evidence that the rate of terrorist

attacks does not change the number of births in the municipality. Moreover, I control directly for terrorist attacks in the pre-pregnancy period, and I find small and statistically insignificant effects for that period, while the main results remain unchanged. Second, I conduct balancing tests that examine whether, across cohorts, variation in terrorist attack exposure is uncorrelated with other time-varying municipality-level shocks. I find no evidence that violence levels are correlated with these shocks. Third, I perform balancing tests to examine whether the variation in violence exposure can explain differences in the characteristics and composition of the cohort. The results show no correlation between violence levels and the number of low birth weights or survival rates.

Finally, given the strong assumption regarding the timing of exposure to terrorist attacks, I perform a nonparametric nearest neighbor match analysis at the municipality level using the average rate of terrorist attacks between 2000 and 2016. This alternative approach does not rely on the staggered design of violence affecting different cohorts or on the assumption that the same violent event yields different outcomes depending on the child's age at the time of exposure. Instead, it assigns all teenagers in a given municipality as either part of the treatment or the control group. It compares their outcomes across municipalities while dealing with the endogeneity of violence. Although direct comparisons of point estimates between the two approaches are not feasible, the direction and statistical significance of the estimates remain consistent, suggesting the robustness of the results.

A key distinction of this work is that I can separate the violence exposure by age intervals (0-5 years and 6-9 years) and by length of exposure. First, the results suggest that violence after 5 years old is driving the results of alcohol consumption decisions, but early childhood exposure significantly affects socioemotional skills, consistent with previous studies (Cunha and Heckman, 2007; Heckman et al., 2013). Second, the results indicate that exposure to violence for five or more years increases the probability of drinking, while

exposure for two or more years increases the total difficulties and hyperactivity scores of the socioemotional test. These results suggest that prolonged exposure during childhood to violence might have long-lasting consequences. Given the literature on the relationship between risky health behaviors, socioemotional skills, and adulthood outcomes, these results underline significant policy concerns about effectively addressing these adverse effects.

This paper contributes to three strands of literature. The first is a set of studies on the impacts of conflict on sociemotional skills and health. While the literature shows that conflict negatively affects health outcomes (Akresh et al., 2012; Minoiu and Shemyakina, 2014; Shemyakina, 2021; Grossman and Khalil, 2022), psychological well-being (Blattman and Annan, 2010), and cognitive and sociemotional skills (Shany, 2023; Jürges et al., 2022) in the short-term (i.e., early childhood and childhood), this study focuses on medium-term (i.e., adolescence) effects similar to León (2012). Notably, it also provides evidence of a new set of outcomes and how adolescents are affected. Moreover, it sheds light on how their decision-making process changes, which could lead to long-term consequences.

The second set of studies is related to the effects of childhood experiences on young adulthood outcomes, mainly focusing on positive experiences and the impacts on education outcomes (Figlio et al., 2014; Bharadwaj et al., 2018). Previous studies have analyzed the effects of preschool programs like Head Start (Ludwig and Miller, 2007; Deming, 2009; Heckman et al., 2013), neighborhood quality (Chetty et al., 2016) or access to safety net programs (Hoynes et al., 2016). This paper contributes to this literature by exploring a prolonged negative experience and analyzing new outcomes. The results suggest that these negative experiences can outweigh the potential positive benefits from other inputs like those previously examined. This study highlights that policies aimed at improving children's adolescent outcomes should consider how to prevent and mitigate the adverse effects of traumatic experiences like violence.

Finally, it contributes to the literature on the impacts of conflict in developing countries like Colombia, where resources to deal with the conflict and its consequences are limited. Previous evidence characterizes the effects of the armed conflict on infant birth weight (Camacho, 2008), child anthropometric measures (Kreif et al., 2022), and school performance and educational attainment (Ramirez Muñoz, 2015; Rodriguez and Sanchez, 2009, 2010). This paper adds to this literature by studying new outcomes (socioemotional skills and risky health behaviors) that are highly important due to their long-term consequences (Heckman et al., 2006; Mezza and Buchinsky, 2021) and are understudied in the context of developing countries.

The remainder of the paper is organized as follows. Background information and a prior literature summary can be found in section 3.2. Section 3.3 provides information on the data, and section 3.4 outlines the empirical strategy. Section 3.5 provides main results, heterogeneity analyses, robustness exercises, and a discussion. I conclude in section 3.6.

3.2 Background and Prior Literature

3.2.1 Armed conflict in Colombia

Colombia has experienced a long-lasting conflict over the past 60 years with two nonstate armed groups - the Revolutionary Armed Forces of Colombia (FARC from the Spanish acronym) and the National Liberation Army (ELN from the Spanish acronym) - founded in the mid-1960s. Initially, the guerrilla groups were located in rural areas where they could maintain territorial control by performing violent actions against the population and infrastructure. However, they tried to expand their territory and control new areas over several decades. In the 1990s, the conflict escalated due to the guerrillas' illegal drug trafficking and consolidation of the United Self-Defense Groups of Colombia

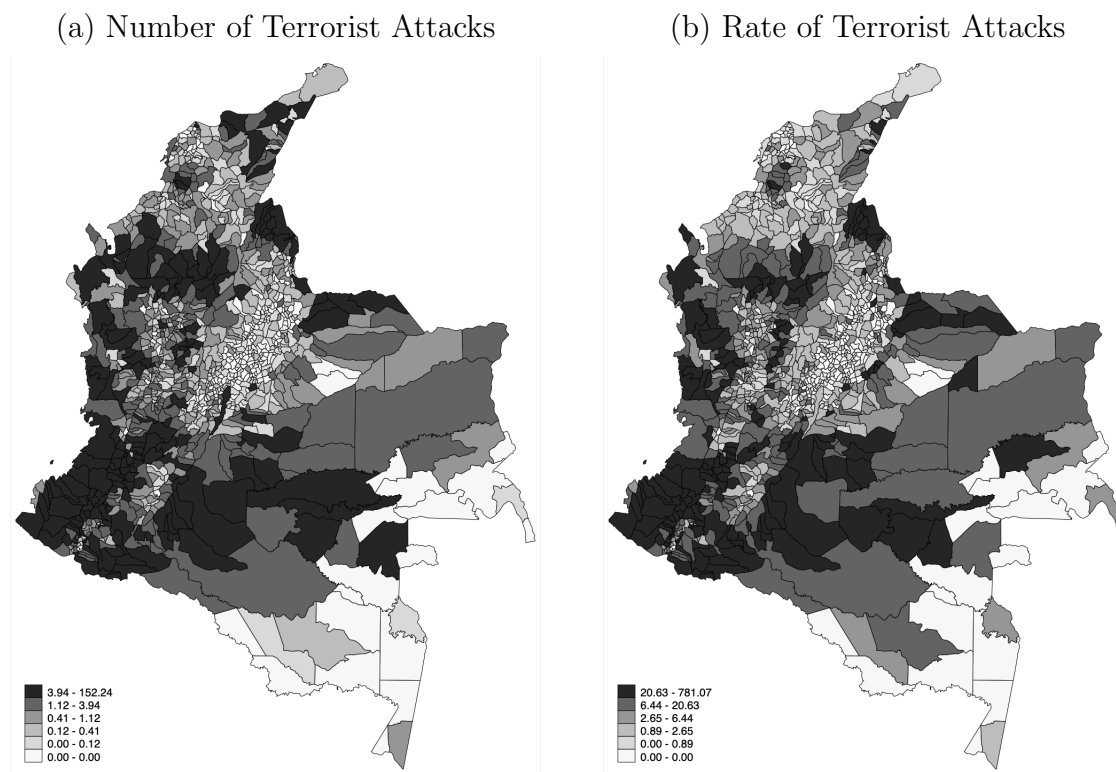
(AUC from the Spanish acronym). The confrontation with these groups dominated the 2002 presidential election in which Álvaro Uribe was elected president. Among his policies was his flagship policy *Política de Seguridad Democrática* (Democratic Security Policy), which included significant military expenditure to combat the armed groups. While the guerrilla withdrew to remote areas between 2000 and 2012, following a failed peace process initiated by President Andrés Pastrana in 1999, terrorist attacks continued to target military and civilian buildings and infrastructure to maintain territorial control. Since the beginning of the conflict, around 9 million people have been registered as victims of the armed conflict.⁹

During the 2000s, violence remained prevalent across Colombia, with significant regional variation. Panel A in Figure 3.1 shows the average number of terrorist attacks per municipality between 2000 and 2016, while Panel B shows the rate of terrorist attacks per 100,000 inhabitants over the same period. Municipalities in black indicate areas more heavily affected by terrorist attacks, primarily in the Pacific region.¹⁰ The Figure highlights two important features of the Colombian conflict. The figure highlights two key features of the Colombian conflict. First, it illustrates the persistently high levels of terrorist attacks during this period and the stark regional differences in the behavior of armed groups, with some regions disproportionately affected. Second, it emphasizes the importance of accounting for population size when analyzing the intensity of exposure to terrorist attacks. For instance, while Bogotá appears darker in Panel A due to the absolute number of attacks, it is shown in light gray in Panel B after adjusting for population. Consequently, this paper uses the rates of terrorist attacks, which account for population size, to better capture the intensity of exposure to terrorist attacks.

⁹Data available from the National Information Network of the Unit for the Victims Assistance and Reparation. Data retrieved from <https://www.unidadvictimas.gov.co/es/registro-unico-de-victimas-ruv/37394> on March 9, 2022.

¹⁰San Andrés and Providencia are not included in the graph as it is located in the insular region

Figure 3.1: Terrorist attacks between 2000-2016



Notes: The figure above shows the geographic variation in terrorist attacks between 2000-2016. Panel (a) estimates the average number of terrorist attacks in the municipality between 2000-2016, while Panel (b) estimates the average rate of terrorist attacks per 100,000 inhabitants in the municipality between 2000-2016. Each panel shows the mean for each of the municipalities in Colombia between 2000-2016, excluding the insular department.

3.2.2 Prior Literature

Reasons to engage in risky health behaviors

The evidence on the impact of adverse childhood experiences (ACE) on risky health behaviors is very limited in the economic literature. The studies have focused mainly on the contemporary factors that induce youth to smoke, drink, or engage in risky sexual health behaviors (Cook and Moore, 2001; Gruber and Zinman, 2001; Levine, 2001; Cawley and Ruhm, 2011). One possible explanation is that youth adopt these behaviors as coping mechanisms in response to distressing life events (Friedman, 2020) or adverse events Shonkoff et al. (2012). This implies that children exposed to violence, for instance, might resort to smoking or alcohol consumption as a coping mechanism to deal with the associated stress.

The evidence suggests that ACEs have long-term negative consequences on adult outcomes. Indeed, exposure to violent situations affects the body and mind's development, particularly during early childhood when the brain's structure and functioning develop continuously (National Scientific Council on the Developing Child, 2015, 2011). This increased sensitivity to change during early childhood is attributed to the brain's plasticity (National Scientific Council on the Developing Child, 2015). Consequently, situations that generate fear and anxiety, such as armed conflict, cause damage to the structure and architecture of the brain (National Scientific Council on the Developing Child, 2010).

Mainly, prolonged exposure to stress affects the prefrontal cortex development, which begins during early childhood and continues throughout adolescence and early adulthood (Arain et al., 2013). This area of the brain plays a critical role in regulating the development of executive functions and responses to emotional stimuli National Scientific Council on the Developing Child (2015, 2010); Shonkoff et al. (2012). The adverse effects of such interference can render adolescents more susceptible to engaging in risky health behaviors,

especially those with a low perceived coping efficacy, as they may struggle to control their impulses (Arain et al., 2013; Asselmann et al., 2016). The intensity of ACE exposure is associated with an elevated risk of behavior problems (Burke et al., 2011). For instance, Dube et al. (2006) find that individuals who experience ACEs are significantly more likely to start drinking alcohol by age 14.

Consequences of violence

Previous research has found evidence of the adverse effects of conflict on health (Akresh et al., 2012; León, 2012; Minoiu and Shemyakina, 2014; Shemyakina, 2021) and psychological well-being (Blattman and Annan, 2010). However, accurately estimating the causal effects of armed conflict is challenging for several reasons. First, identifying victims of violence can be difficult due to confidentiality issues. For this reason, self-reports in surveys are often used, leading to potential non-random measurement error, recall bias, or reluctance to share experiences to avoid revictimization (Ellsberg et al., 2001). For this reason, I use the violence levels at the municipality level reported by the National Information Network. Second, violent victimization is not random, as armed groups seek control over strategic regions to finance their activities. Consequently, areas where armed groups can gain political power or control of rent extraction are more likely to experience violent events (Centro Nacional de Memoria Histórica, 2013; Fernández et al., 2014). Therefore, isolating the causal effect of violent crimes from other influencing factors at the individual, community, or municipal level can be challenging. To address this, rather than comparing different municipalities, I analyze cohorts within the same municipality that have been exposed to varying levels of violence.

The economics literature has focused on two main topics: (i) estimating the causal effects of conflict on health, education, and psychological well-being in the short-term, and (ii) estimating the long-term effects of “positive” interventions like preschool programs.

More recently, the literature has started to analyze the impact of “negative” environments on medium-term and long-term outcomes. For instance, Jürges et al. (2022) find that parental exposure to conflict has adverse long-term effects on educational attainment, with non-cognitive skills playing a key role in these outcomes. However, the impact of children’s direct exposure to conflict during childhood on their adolescent outcomes remains unknown. Moreover, Manian (2021) shows that conflict leads to an increase in risky sexual behavior (such as unprotected sex) among adults and that behavioral responses can mitigate the negative effects of conflict by enhancing risk aversion. Finally, the empirical literature on long-term effects focuses on chronic diseases (Allais et al., 2021; Di Maio and Sciabolazza, 2021).

3.3 Data

For this study, I use two sources of information: (i) the 2010 and 2016 Waves of the Colombian Longitudinal Survey (ELCA) and (ii) the Municipal Panel of the Center for Studies on Economic Development (CEDE).

3.3.1 Colombian Longitudinal Survey

The ELCA is a nationally representative longitudinal survey that follows households (strata 1 to 4) since 2010 and their members residing in urban areas across five regions: Atlantic, Eastern, Central, Pacific, and Bogotá. It collects information on household composition, expenses, income, and caretaker characteristics. The analysis sample includes 870 adolescents aged 13 to 16 residing in 832 households. The sample includes children among the specified age range for two reasons. Firstly, the evidence indicates that the typical onset of alcohol consumption among Colombian children is around age 12 (de la

Espriella Guerrero et al., 2016).¹¹ Secondly, the survey design excludes any adolescent older than 16 from the comprehensive questionnaire, resulting in a lack of information on risky behaviors and socioemotional skills for individuals in this age bracket.¹²

This study uses information collected from the survey on socioemotional skills and risky health behaviors. First, the survey includes information on socioemotional skills, assessed using the Spanish-approved version of the Strengths and Difficulties Questionnaire (SDQ), a brief behavioral screening questionnaire for children up to 17 years old. The SDQ measures the total number of difficulties, allowing analysis across four distinct components: Emotional symptoms, conduct problems, hyperactivity/inattention, and peer relationship problems.¹³ The SDQ operates on a scale ranging from 0 to 40 points, with each component scored out of 10 points. Higher scores indicate a higher degree of problems. The total score and each component are used as the outcomes in this paper.

Second, the risky health behaviors information includes questions on alcohol consumption, personal smoking habits, and the smoking habits of friends. These questions are self-reported. To mitigate potential measurement errors during the survey, parents are not allowed in the room when the respondents provide their answers to these questions. The survey team took this precaution because teenagers tend to offer less accurate information when their parents are present.¹⁴ This paper uses 4 outcome variables based on this data: (i) binary indicator of whether the respondent has tried alcohol, (ii) age of first drink, (iii) binary indicator of whether the respondent smokes, and (iv) binary indicator of whether the respondent's friends smoke.

¹¹Consistent with this, only 23% of the children 12 or below report having tried alcohol compared to 64% of those older than 13.

¹²See Online Appendix C.2 for a discussion of how the sample was selected. Also, for more information on the questionnaires or data access, see <https://encuestalongitudinal.uniandes.edu.co>

¹³It is also possible to diagnose externalizing problems (conduct and hyperactivity scores) and internalizing problems (emotional and peers score)

¹⁴When parents cannot leave the room or refuse to, the enumerator gives the respondent a card to signal their answers to each question without speaking aloud.

Table 3.1: Summary Statistics

	Mean	SD	Median
A. Adolescents characteristics			
Sex (1 = Male)	0.50	0.50	0.00
Age in years	14.04	0.88	14.00
Years of education	7.02	1.59	7.00
B. Household characteristics			
Two parent household	0.57	0.49	1.00
Household size	5.43	2.49	5.00
Total annual expenditure (in millions)	15.61	23.96	10.89
C. Caretaker characteristics			
Age	41.63	8.38	40.00
Years of education	9.06	4.39	11.00
Participation in social organizations (Number)	0.21	0.52	0.00
Leader in a social organizations (Number)	0.08	0.34	0.00
D. Municipality characteristics			
Rurality index	0.21	0.20	0.14
Distance to department capital	42.27	49.61	26.55
Unsatisfied Basic Needs (Urban in 2005)	21.63	15.81	17.24
Population in 2010 (in thousands)	406.44	1,033.10	111.20
Rate of terrorist attacks 0-9	3.68	8.39	0.98
Homicide rate 0-9	105.65	134.08	62.89

Notes: The table above reports summary statistics of the municipalities and adolescents in the sample. It includes adolescents's characteristics and those of their households and caretakers. The table only includes the municipalities in the sample of analysis, not all of the country. Adolescent variables are from the 2016 survey, while household and caretaker characteristics are from the baseline (2010).

Table 3.1, Panels A, B, and C provide summary statistics for the adolescents in the sample and their households. Notably, there is little variation in adolescents' characteristics except in their education levels. In contrast, household and caretaker characteristics exhibit more variation, particularly concerning variables such as household size, expenditure, and

caretakers' education level. Additionally, there are no significant differences across cohorts in adolescent (except age), household, and caretaker characteristics, which are necessary to ensure the cohorts are comparable (see Online Appendix Table C.1). Similarly, no significant differences in characteristics are observed across regions (see Online Appendix Table C.2).

3.3.2 Municipal Panel

The Municipal Panel of the Center for Studies on Economic Development (CEDE) provides data on various indicators categorized by subject (e.g., General Characteristics, Conflict, and Violence). The violence data comes from official reports of the National Information Network between 2000 and 2012. The data includes several forms of violence, including terrorist attacks, the primary measure of exposure to conflict in this study. Based on this data, two measures of terrorist attack exposure are constructed. First, the average rate of terrorist attacks per 100,000 inhabitants in the municipality when the children were between 0 and 9 years old. This measure accounts for population size and considers that larger municipalities may experience more terrorist attacks. Still, the adolescent may not have been exposed. I provide specification checks using only the smaller municipalities where children are more likely to be exposed to a municipality-level shock. Second, the number of years that an adolescent was exposed to at least one terrorist attack during her first nine years.

Additionally, supplementary data regarding the municipality includes the rurality index and the Unsatisfied Basic Needs (UBN) of the urban areas within the municipality.¹⁵ These variables were obtained through the National Administrative Department of Statistics (DANE), the Geographic Institute Agustín Codazzi (IGAC), or calculated by the CEDE

¹⁵The UBN is an index that contains four dimensions: housing and related indicators, water and sewage indicators, education and associated indicators, and income-related indicators

using primary information.

The ELCA conducted interviews with adolescents across different municipalities in Colombia, each facing varying levels of conflict. Table 3.1, Panel D provides summary statistics for the municipalities within the sample. The table underscores the predominantly urban nature of these municipalities, as evidenced by the low rurality index. This is consistent with the sample selection process, as the ELCA sample exclusively covers households living in urban areas. Additionally, the municipalities are close to their respective department capitals and display lower levels of Unsatisfied Basic Needs. The table indicates that the sample is skewed towards more urbanized areas, emphasizing that the study results apply solely to urban households in Colombia. Regarding violence levels, homicides are the most prevalent form, with terrorist attacks also prevalent, albeit with variations across different municipalities.

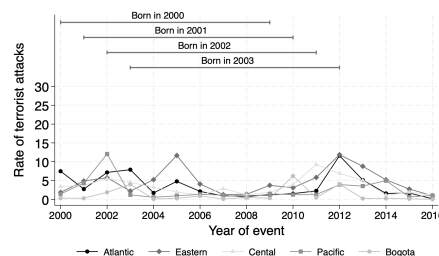
Due to data restrictions, the specific municipalities included in the sample are not disclosed. Therefore, the results are presented at the regional level, with the distributions of raw variation in violence levels shown where possible. Panel A in Figure 3.2 shows the average rate of terrorist attacks per 100,000 inhabitants for municipalities in the sample across five regions of the country by year. This panel highlights two significant findings. First, as expected, there is substantial variation in violence levels across regions. Second, no national trend suggests that the conflict disproportionately impacted any specific cohort compared to others.¹⁶ Furthermore, Panel B illustrates the distribution of the average rate of terrorist attacks by municipality, while Panel C shows the within-municipality standard deviation of the rate. These panels emphasize that, while terrorist attacks are a

¹⁶The absence of differences across cohorts could be explained by differences in levels. Online Appendix Figure C.1 depicts the average rate of terrorist attacks per 100,000 inhabitants experienced by each cohort in each region without adjusting for differences in levels across municipalities, while Online Appendix Figure C.2 estimates the average of the demeaned rate of terrorist attacks, accounting for these level differences. Both figures suggest that, even after adjusting for these differences, violence levels still vary across cohorts.

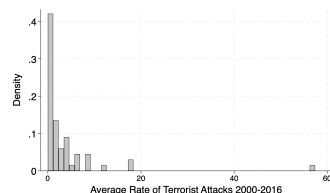
relatively rare form of violence, most municipalities in the sample experienced at least some attacks. Additionally, they highlight the significant variation in the frequency of these attacks over time. Both panels also indicate that one municipality experienced a substantially higher number of terrorist attacks. In the results section, I will present a sensitivity analysis demonstrating that the inclusion of this municipality does not drive the findings.

Figure 3.2: Average Rate of terrorist attacks between 0-9 years old, split by region

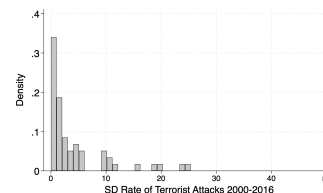
(a) Average rate of terrorist attacks by region and year



(b) Distribution of the Average rate of terrorist attacks by municipality



(c) Distribution of the Within-Municipality standard deviation of the rate of terrorist attacks

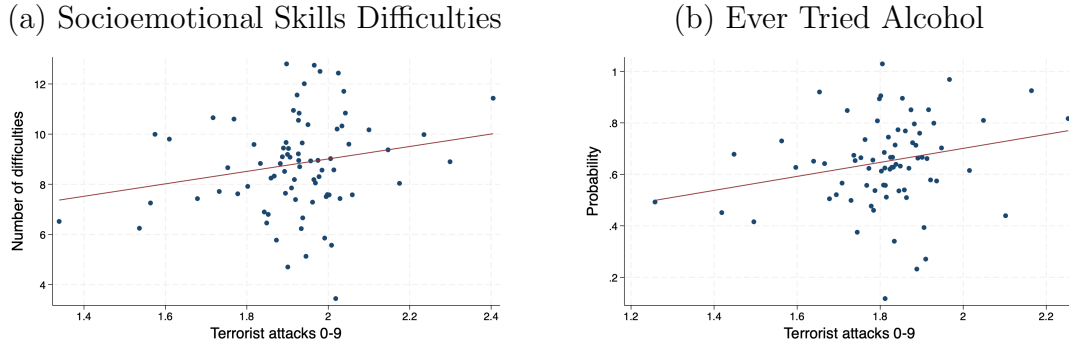


Notes: The figure above shows the variation in the rate of terrorist attacks between 2000-2016 for the municipalities in the sample. Due to data restrictions, the municipalities in the sample are considered restricted access data. Panel (a) estimates the average per region and year. It also shows the years used to create the rate of terrorist attacks between 0-9 years old for each cohort. Panel (b) plots the distribution of the average rate of terrorist attacks between 2000-2016 for each municipality, while panel (c) shows the distribution of within-municipality standard deviation in the rate of terrorist attacks between 2000-2016. The regions follow the same division as the Colombian Longitudinal Survey (ELCA).

3.3.3 Preliminary checks

Before estimating the impacts of the rate of terrorist attacks an adolescent was exposed to between 0 and 9 years old on socioemotional skills and risky health behaviors, I provide some preliminary descriptive evidence. Figure 3.3 illustrates the relationship between the rate of terrorist attacks an adolescent was exposed to between 0 and 9 years old and the two main outcomes. The figure shows the residualized means of the rate of terrorist attacks and each outcome, conditional on municipality and cohort fixed effects. Specifically, it compares adolescents living in the same area with differential exposure to municipality terrorist attacks during childhood while accounting for national shocks that occurred in a specific year. Panel A suggests that adolescents exposed to a higher frequency of terrorist attacks in their first nine years of life tend to obtain higher scores in the total difficulties score, indicative of a potential depletion in their socioemotional skills. Panel B demonstrates a similar pattern for alcohol consumption, as the percentage of adolescents who report having started drinking alcohol is positively correlated with exposure to terrorist attacks during childhood.

Figure 3.3: Correlation between Terrorist Attacks and Socioemotional Skills and Alcohol Consumption



Notes: The figure above shows the relationship between the rate of terrorist attacks a child was exposed to between 0 and 9 years old and the two main outcomes. The figure shows residualized means, conditional on municipality and cohort fixed effects. Socioemotional skills difficulties are measured as the total number of difficulties from the Strengths and Difficulties Questionnaire (SDQ) test score, and a higher score reflects that the child has more problems. Ever tried alcohol is a binary indicator that measures whether the children had tried alcohol before the time of the survey (2016). .

3.4 Empirical Strategy

To analyze the impact of terrorist attack exposure on socioemotional skills and risky health behaviors, I combine the ELCA with the Municipal panel. The estimations do not leverage the panel structure of the ELCA, given that the outcomes are solely available in the 2016 wave. Instead, the panel structure is exclusively used to control for baseline characteristics and to create measures of terrorist attack exposure during childhood (from 0 to 9 years old for each child), based on the adolescents' municipality of residence in 2010 and their birth year. Terrorist attack exposure variation stems from two different sources: the individual's birth year and the municipality of residence in 2010. Panel A

in Figure 3.2 displays the variation in the levels of violence experienced by each cohort. Most of the variation comes from the first and last year, as the four cohorts are staggered, and the adolescents were born only one year apart.

To estimate the average causal effect of terrorist attack exposure on socioemotional skills and risky health behaviors, I estimate the following baseline two-way fixed effects specification, using OLS:

$$y_{imt} = \beta_0 + \beta_1 V_{mt}^{0-9} + \beta_2 V_{mt}^{10+} + \mathbf{X}_{imt}\beta_3 + \mathbf{Z}_{imt}\beta_4 + \mathbf{C}_{imt}\beta_5 + \mu_m + \delta_t + \varepsilon_{imt} \quad (7)$$

where y_{imt} is one of the outcomes available in the ELCA for adolescent i in municipality m from cohort t . V_{mt}^{0-9} represents the average rate of terrorist attacks per 100,000 inhabitants in the municipality m experienced by adolescents in cohort t during their childhood. V_{mt}^{10+} represents the average rate of terrorist attacks per 100,000 inhabitants in the municipality m experienced by adolescents in cohort t experienced after 10 years old. $\mathbf{X}_{imt}$ is a vector of individual-level controls such as gender, educational level, order within the children in the household¹⁷; $\mathbf{Z}_{imt}$ is a vector of household-level controls including the logarithm of consumption, whether the household is a two-parent household, and household size; $\mathbf{C}_{imt}$ is a vector of caretaker-level controls including age and years of education. μ_m are municipal fixed effects, and δ_t are cohort fixed effects. The standard errors are clustered at the municipality level to account for serial correlation within each municipality.

Comparing adolescents across municipalities would be inappropriate due to concerns surrounding the non-random nature of municipal violence, as explained in section 3.2.2. The two-way fixed effects model presented above is used to address this issue. Including municipality-specific fixed effects μ_m accounts for systematic differences between municipalities and their prevalence of armed conflict. Consequently, these fixed effects ensure

¹⁷This variable is an ordering of the children in the household from eldest to youngest. This means that the oldest child is assigned one, and so with every child until reaching the last one

that the identification stems from within-municipality variations. Simultaneously, cohort-specific fixed effects account for shocks occurring in a specific year (i.e., national-level shocks). Hence, both sets of fixed effects are included to identify the causal effects of an adolescent's exposure to terrorist attacks in a given municipality compared to other adolescents in other cohorts in the same municipality that experienced different levels of terrorist attacks. With this specification, the parameter of interest, β_1 , measures the impact on socioemotional skills and risky health behaviors of the average exposure to terrorist attacks during childhood.

However, given that the two-way fixed effects model relies on a strong assumption, I perform a nonparametric nearest neighbor match analysis at the municipality level using the average terrorist attacks between 2003 and 2013. For this, municipalities are assigned to the treated group if the average rate of terrorist attacks between 2000 and 2013 is greater than the median and to the control group otherwise. This approach does not rely on the staggered design of how violence affects different cohorts, and it does not rely on the assumption that the same violent event leads to different results depending on the child's age at the moment of exposure. Instead, this approach assigns all teenagers in a given municipality to the treatment or control group and compares teenagers across municipalities while dealing with the endogeneity of the violence. In particular, it matches individuals across municipalities based on individual, household, and municipality characteristics.

Identification

The goal of estimating Equation 7 is that β_1 will identify the causal effect of terrorist attacks. The identification strategy relies on the premise that while parents may base decisions on the general characteristics of a municipality, they do not do so based on the specific characteristics of their child's cohort within the municipality, including the

level of violence their child might experience. Conditional on municipality and cohort fixed effects, within-municipality variation in violence levels is not correlated with other factors (observed and unobserved) affecting the studied outcomes. One feature of terrorist attacks is their unpredictable nature, even for military forces, and their inherent element of surprise, which provides ex-ante support for identification. Indeed, these events catch the population off guard and should have substantial effects.

Table 3.2: Variation in Rate of Terrorist Attacks After Removing Cohort and Municipality Fixed Effects

	Mean	SD	Median	Min	Max
A. Raw variables					
Rate of terrorist attacks 0-9	3.42	8.73	1.35	0.00	72.00
B. Residuals after removing cohort and municipality FE					
Residuals of Rate of terrorist attacks 0-9	0.02	1.33	0.02	-11.89	7.13

Notes: The table above reports the raw variation in the average rate of terrorist attacks a child was exposed to between 0 and 9 years old and the variation that is left after removing cohort and municipality fixed effects. The table only includes the municipalities in the sample of analysis. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

The success of the main specification presented in Equation 7 depends on three key factors. Second, to obtain precise estimates, there must be sufficient variation in the exposure to terrorist attacks after controlling for cohort and municipality fixed effects. Table 3.2 examines the extent of variation in exposure to terrorist attacks before and after accounting for these fixed effects. As expected, removing the fixed effects reduces the

standard deviation in the rate of terrorist attacks by over 80 percent. Therefore, the effect estimates in this paper are based on small changes in exposure. Despite this reduction, there remains sufficient variation to estimate the effects of these small changes in exposure with relative precision, as the results presented later in the paper demonstrate.

Second, in order to make causal interpretations of the estimates, within-municipality variation in terrorist attacks must be uncorrelated with differences in time-varying municipality-level shocks. If these variables are uncorrelated, the analysis supports the premise that municipality and cohort fixed effects capture any systematic selection on observables. For this reason, I regress the average municipality-level characteristics on the average rate of terrorist attacks experienced by each cohort during childhood, controlling for cohort fixed effects and municipality fixed effects. Panel A in Table 3.3 shows that the variation in the exposure to terrorist attacks is not correlated with the variation in the characteristics at the municipality level.

Third, the within-municipality variation in terrorist attacks must be uncorrelated with differences in demographic characteristics across cohorts. For this reason, I regress the average number of births, number of low birth weights, and number of deaths under age four on the average rate of terrorist attacks experienced by each cohort during childhood, controlling for cohort fixed effects and municipality fixed effects. Panel B in Table 3.3 shows that the composition and characteristics of the cohorts are not correlated with the variation in the exposure to terrorist attacks.

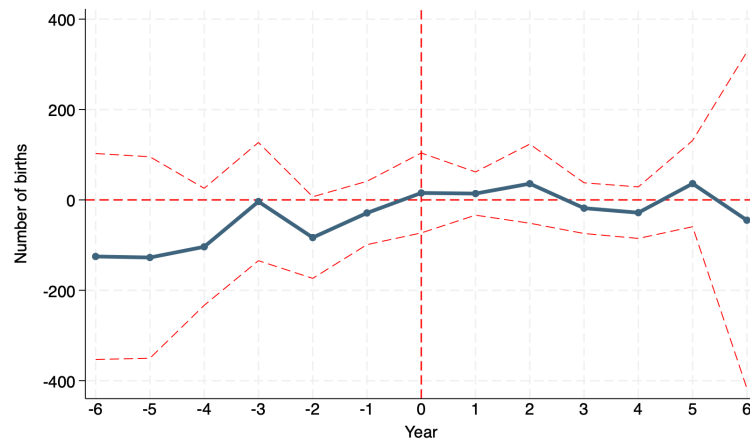
Table 3.3: Balancing Tests for Municipality Shocks and Characteristics

	Rate of Terrorist Attacks
A. Municipality-Level Characteristics	
Rurality index	-0.000 (0.000)
Gross Domestic Product (billions)	-0.066 (0.704)
Income (billions)	0.134 (2.647)
Tax Income (billions)	0.191 (1.202)
Expenditure (billions)	0.061 (2.600)
Deficit (billions)	0.073 (0.100)
Central Government Transfers (billions)	0.037 (0.544)
Central Government Transfers for Education (billions)	0.050 (0.401)
Central Government Transfers for Health (billions)	-0.016 (0.087)
Number of Primary school age children	8.243 (14.235)
Number of Secondary school age children	13.359 (24.960)
Number of schools	0.123 (0.166)
Number of teachers	1.425 (2.872)
Number of students	37.572 (40.381)
B. Cohort Characteristics	
Number of births	0.819 (6.241)
Number of low birth weights	0.269 (0.569)
Number of deaths of children under the age of four	-0.005 (0.064)

Notes: The table above reports balancing tests for municipality and cohort characteristics. The figures in each row are coefficients from regressions that include, in addition to the rate of terrorist attacks, controls for municipality and cohort fixed effects. All variables are measured using information from the Municipal Panel of the CEDE. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Figures in parentheses are standard errors corrected for heteroskedasticity and clustered at the municipality level.

However, recognizing the potential threat to identification, I address it in several ways. First, the timing of birth could potentially be endogenous, driven by factors such as parents in more violent municipalities delaying pregnancy or being able to predict periods of increased terrorist attacks, influencing their fertility decisions. In Figure 3.4, I show the impact of the rate of terrorist attacks six years prior and after on the number of births in a given year. For instance, the number of children born in 2000 is regressed on the rate of terrorist attacks each year between 1994 and 2006. The findings indicate that prior violence in the municipality does not influence the number of births, suggesting that parents are not adjusting the timing of deliveries in response to violence levels.¹⁸

Figure 3.4: Impact of Terrorist Attacks on Number of Births



Notes: The figure above reports estimates from regressions of the rate of terrorist attacks per 100,000 inhabitants 7 years prior and 7 years after on the number of births in a given cohort. The solid line represents the point estimate, and the dashed lines represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

¹⁸Violence could also impact the survival rate of children born in each cohort. Panel A in Online Appendix Figure C.3 shows the survival rate of children born in each cohort over time. Panel B shows no significant correlation between the survival rate and the rate of terrorist attacks during childhood.

Second, I test the sensitivity of the estimates to different ways of controlling for differences between cohorts. This involves incorporating either cohort-by-region fixed effects or neighborhood fixed effects. The cohort-by-region specification leverages within-region variation in the timing of terrorist attacks across municipalities, accounting for region-level shocks or policies influencing a cohort's outcomes. The neighborhood fixed effects specification considers shocks experienced by adolescents in the same neighborhood that might impact their outcomes. In both sets of specifications, the main estimates are similar. Furthermore, I explore whether excluding the cohort fixed effects and only controlling for age changes the results. However, the results do not show significant changes.

Third, sensitivity tests are conducted to address any concerns regarding omitted factors influencing the results. Specifically, I estimate two alternate specifications: i) limiting the sample to municipalities with a predicted propensity for terrorist attacks within the range of $[0.1, 0.9]$, based on their time-invariant characteristics; ii) excluding the four largest cities in the country from the main specification. These results are also similar to those obtained from the main specification.

Fourth, the rate of terrorist attacks is constructed using the municipality of residency in 2010. However, parents may adapt to violence levels by migrating from more violent to less violent areas. For this reason, I perform a set of analyses where terrorist attack exposure is constructed using the birthplace. As will be discussed in the next section, the results are robust to this adjustment.

Lastly, given that the sample only includes four cohorts, it implicitly assumes that the timing of exposure to the same violent event yields different results contingent on the child's age at the moment of exposure. Although direct testing of this assumption is not feasible with the available data, as discussed in section 3.2.2, the brain's structure and functioning develop continuously throughout early childhood. Due to its plasticity, the

brain is much more sensitive to changes during this period (National Scientific Council on the Developing Child, 2015). Previous research on safety net programs exposure during childhood has assumed and shown that the impact of these programs changes depending on the age of exposure (Hoynes et al., 2016; Bailey et al., 2024), which is also consistent with previous findings in the early childhood literature (Heckman and Rubinstein, 2001). Consequently, the timing of exposure is likely to be a crucial factor, with potentially more severe consequences for younger children actively undergoing developmental processes and possessing more plastic brains (National Scientific Council on the Developing Child, 2010).

3.5 Results

In this section, the estimated causal effects of exposure to terrorist attacks during childhood on socioemotional skills and risky health behaviors are reported using OLS. Unless otherwise noted, all analyses in this paper utilize the preferred specification given by Equation 1.

3.5.1 Main Results

Table 3.4 reports that exposure to terrorist attacks during childhood significantly depletes socioemotional skills among Colombian adolescents living in urban areas. All the columns show the baseline specification from Equation 1, which includes cohort and municipality fixed effects. Odd columns do not include controls, while even columns do. Panel A presents the estimates obtained from Equation 1. Column 2 in this panel shows that, on average, exposure to one additional terrorist attack during childhood leads to a 0.046 SD (0.25 points) increase in the total socioemotional difficulties score of the Strengths and Difficulties Questionnaire. In contrast, the estimate for terrorism exposure

during early adolescence is not statistically significant. Notably, this result is mainly driven by the hyperactivity domain score, which increases by 0.064 SD (0.14 points) as shown in Column 4. These effects are smaller than those previously reported in the literature Jürges et al. (2022). However, this paper does not analyze direct victimization as in previous papers. Therefore, the estimates should be smaller. No significant effects are observed on the other dimensions of the SDQ score. These findings are consistent with the notion that terrorist attack exposure in a sensitive period, like childhood, interferes with the proper functioning and development of the brain. Hence, the exposure leads to a depletion of skills, including socioemotional. Panel B in Figure 3.4 presents the nearest neighbor match design results and also suggests a depletion in socioemotional skills due to exposure to violence during childhood. Although comparing the magnitude of the results between the fixed effects approach and the nearest neighbor match is design is not plausible, the direction of the effects is the same and are statistically significant.

Table 3.4: Effects of Terrorist Attacks on Socioemotional Skills

	Total		Score components							
	Score		Hyperactivity		Emotional		Peers problems		Conduct	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Within-Municipality										
Cross-Cohort Variation										
Terrorist Attacks 0-9 years old	0.046*	0.046**	0.046**	0.056**	0.042**	0.028	0.040	0.043	-0.023	-0.023
	(0.024)	(0.023)	(0.023)	(0.024)	(0.021)	(0.018)	(0.044)	(0.045)	(0.031)	(0.030)
Terrorist Attacks after 10 years old	-0.011	-0.008	-0.043	-0.013	0.040	0.012	0.023	0.026	-0.048	-0.053
	(0.035)	(0.036)	(0.032)	(0.038)	(0.038)	(0.037)	(0.044)	(0.048)	(0.043)	(0.041)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B. Nearest Neighbor Match										
Terrorist Attacks 0-9 years old years old > Median		0.140***		0.063***		0.182***		0.018***		0.114***
		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
Municipality FE	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Mean of Dependent Variable	8.811	8.811	3.355	3.355	1.939	1.939	1.300	1.300	2.227	2.227
Observations	856	856	856	856	856	856	856	856	856	856
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Panel A in the table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Socioemotional Skills Scores. Panel B reports estimates from the nearest matching neighbor estimations. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table 3.5: Effects of Terrorist Attacks on Alcohol Consumption

	Ever tried alcohol		Age of first drink	
	(1)	(2)	(3)	(4)
A. Within-Municipality				
Cross-Cohort Variation				
Terrorist Attacks 0-9 years old	0.040*** (0.012)	0.042*** (0.012)	-0.070 (0.146)	-0.062 (0.125)
Terrorist Attacks after 10 years old	0.017 (0.017)	0.023 (0.016)	0.163 (0.137)	0.159 (0.133)
Municipality FE	Yes	Yes	Yes	Yes
B. Nearest Neighbor Match				
Terrorist Attacks 2000-2016 > Median		0.006*** (0.001)		-0.244*** (0.002)
Municipality FE	Yes	No	Yes	No
Mean of Dependent Variable	0.643	0.643	12.84	12.84
Observations	870	870	495	495
Controls	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes

Notes: Panel A in the table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Alcohol Consumption. Panel B reports estimates from the nearest matching neighbor estimations. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Column 2 in Table 3.5 shows that, on average, the probability of drinking increases by 4.0 percentage points with one additional terrorist attack during the first years of life. Recognizing that exposure to terrorist attacks may alter the age at which individuals

first consume alcohol, Column 3 indicates that there are no significant reductions with exposure to one additional terrorist attack during childhood. However, the results are noisy and do not indicate that the effects are precisely estimated. Additionally, Online Appendix Table C.3 shows no significant effects on smoking behavior, both concerning the teenagers' decisions and their friends when using the two-way fixed effects. In contrast, the nearest neighbor match design suggests that both probabilities increase with violence exposure. These findings are consistent with the notion that terrorist attacks instill fear and anxiety in the people living in the municipality, prompting them to engage in risky health behaviors. As previously mentioned, the direction of the effects of both estimation approaches is the same and are statistically significant.

As explained in section 3.3.2, there is one municipality that experienced a substantially higher number of terrorist attacks. Therefore, to ensure that the results are not driven by a single municipality, Appendix Figures C.4 and C.5 show the leave-one-out coefficient estimates for all total difficulties score and the probability of drinking. In particular, each regression is estimated omitting one municipality in each iteration—all of which demonstrate similar findings to the main results. The municipality with the highest number of terrorist attacks corresponds to the last estimate on the right side of each panel.

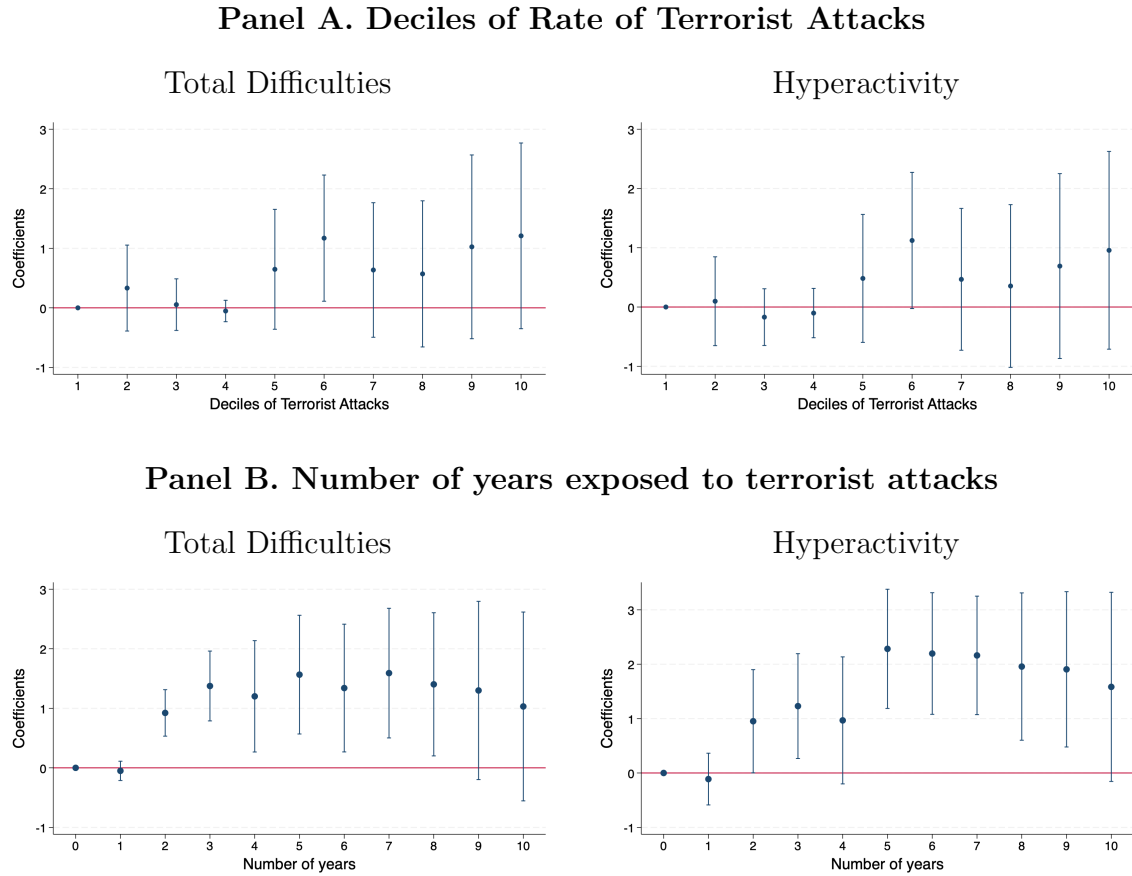
One concern is whether the choice of fixed effects in the main specification drives the results. Section 3.5.4 discusses the results of the specification checks and tests performed in more detail. Nevertheless, if the specification drove the results, alternative forms of violence should yield similar results. To explore this, the average rate of homicides per 100,000 inhabitants during childhood for each cohort was constructed. As shown in Online Appendix Table C.4, there are no significant effects of exposure to homicides during childhood on socioemotional skills. Similarly, Online Appendix Table C.5 shows no significant effects on both alcohol-related outcomes. These findings suggest that the

specification is not driving the main results. There are three plausible explanations. First, terrorist attacks are more likely to create panic and fear at the municipal level, even among those not directly affected, as they often target civilian buildings, general infrastructure, and military forces. Second, while homicides and other forms of violence are more common, they typically target specific individuals or small groups. Consequently, they are less likely to directly or indirectly affect the children not involved. Third, the surprising nature and unpredictability of terrorist attacks make people more likely to be negatively affected by them as they are caught off guard.

3.5.2 Non-Linear Effects

This section explores the non-linearities in the effects of terrorist attacks on socioemotional skills and alcohol consumption. Two different measures of non-linearity are constructed. First, the rate of terrorist attacks is binned into deciles. This measure assesses whether the impacts of exposure are stronger with increased exposure levels. Second, for each cohort, the number of years during childhood that the adolescent was exposed to terrorist attacks is constructed. This measure assesses the long-term exposure, which may be more closely related to the fear and anxiety generated by conflict discussed in section 3.2. Specifically, the longer children are exposed, the more likely they are to develop increased fear and anxiety about their future.

Figure 3.5: Non-Linear Impacts of Terrorist Attacks on Socioemotional Skills



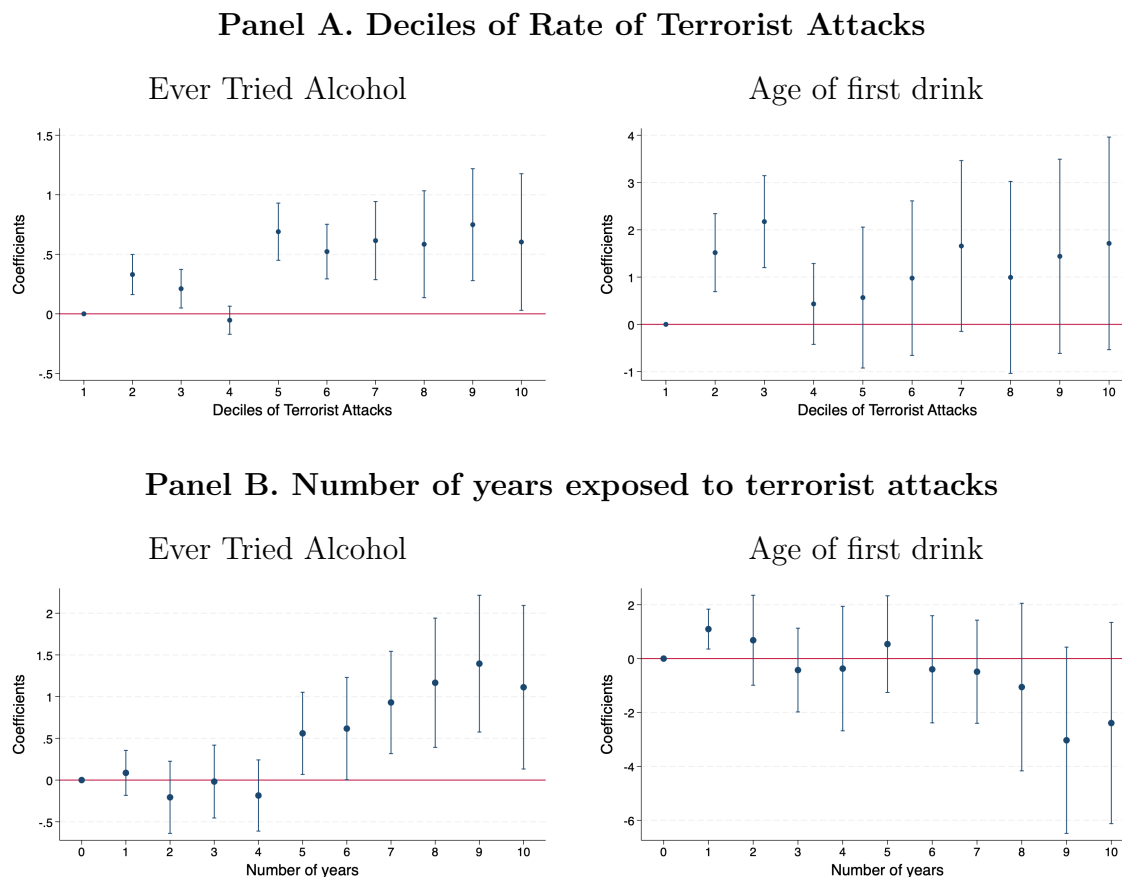
Notes: Panel A in the figure above reports estimates from regressions of the deciles of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Socioemotional Skills Scores. The first decile is the omitted category. Panel B in the figure above reports estimates from regressions of the number of years with strictly greater than 0 terrorist attacks exposure between 0 and 9 years old on Socioemotional Skills Scores. 0 years of exposure is the omitted category. All models include controls for children’s biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. Results are weighted by children’s ELCA weights. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Figure 3.5 presents the non-linearities in the socioemotional skills scores. Panel A shows that the coefficient estimates for higher deciles seem larger than those for the lowest deciles. However, none of these results are statistically significant. Therefore, it is unclear whether the increased exposure severity depletes socioemotional skills. Panel B presents the results for the years exposed to terrorist attacks. In this regression, the number of attacks over the year changes; any year with any attack is included in the estimation. The magnitude of the effects increases with the duration of exposure. Specifically, exposure for two or more years increases the total difficulties and hyperactivity scores. Moreover, the effects on hyperactivity imply that exposure for 5 or more years also depletes socioemotional skills.¹⁹ These results highlight that most of the effect on socioemotional skills is through hyperactivity.

Figure 3.6 presents the non-linearities in the alcohol consumption outcomes. Panel A shows that the magnitude of the effects on the probability of drinking increases with the severity of exposure. However, the results concerning the age of the first drink are notably noisier, likely attributed to the smaller sample size, making it unclear how the severity of exposure influences this outcome. Panel B presents the results for the years exposed to terrorist attacks. The magnitude of the effects increases with the duration of exposure. Specifically, exposure for five or more years increases the probability of drinking. These results underscore the significance of childhood and how exposure for an extended period might have long-term consequences.

¹⁹Online Appendix Figure C.7 presents the results for the other components.

Figure 3.6: Non-Linear Impacts of Terrorist Attacks on Alcohol Consumption



Notes: Panel A in the figure above reports estimates from regressions of the deciles of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Alcohol Consumption Outcomes. The first decile is the omitted category. Panel B in the figure above reports estimates from regressions of the number of years with strictly greater than 0 terrorist attacks exposure between 0 and 9 years old on Alcohol Consumption Outcomes. 0 years of exposure is the omitted category. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. Results are weighted by children's ELCA weights. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Overall, the results suggest that both severity and duration of exposure induce more substantial adverse effects in adolescents affected. Given the existing literature on the impacts of risky health behaviors and socioemotional skills on adulthood, these results raise essential policy concerns about the future consequences on the lives of affected children as they transition into young adulthood and adulthood.

3.5.3 Results by periods of life

The results concerning the duration of exposure highlight that when exposure starts in early childhood, the impacts are more pronounced. Consequently, this section explores whether shocks experienced in early childhood (0-5) shocks have stronger effects than those experienced later in childhood (6-9).²⁰ Parsing the impacts in this way allows testing if the primary mechanism through which the violence influences the outcomes is by worsening the home environment or noncognitive skills of young children due to the fear generated. If this hypothesis holds, the impacts of terrorist attacks should be larger during the earlier period (0-5 years old). As in the main specification, the estimates are interpreted as the impact of the exposure to one additional terrorist attack on the main outcomes but separately for each period of life.

Table 3.6 provides suggestive evidence²¹ that exposure during early childhood depletes socioemotional results. Specifically, on average, experiencing one more terrorist attack increases the total difficulties score by 0.058 (Column 2) and the hyperactivity score by 0.044 (Column 4).²² These results are consistent with the evidence that early childhood shocks have long-term consequences on the socioemotional skills development process

²⁰Bastian and Micheltore (2018) and Bailey et al. (2020) use a similar approach to analyze the effect of the Earned Income Tax Credit (EITC) and Food stamps

²¹As explained in section 3.4, I assume that the timing of exposure to the same violent event leads to different results and has differential importance for the children. This assumption is even stronger in this analysis.

²²Online Appendix Table C.6 presents the results for the other components.

(Cunha and Heckman, 2007; Ludwig and Miller, 2007; Deming, 2009; Almond et al., 2018). More importantly, the results underscore the importance of implementing programs and policies in conflict-affected settings to mitigate the negative effects on children exposed early in life (Sánchez-Ariza et al., 2023).

Table 3.6: Effects of Terrorist Attacks on Socioemotional Skills by Childhood periods

	Total Score		Hyperactivity	
	(1)	(2)	(3)	(4)
Terrorist Attacks 0-5 years old	0.064*** (0.019)	0.058*** (0.018)	0.036 (0.025)	0.044** (0.021)
Terrorist Attacks 6-9 years old	-0.054** (0.022)	-0.043* (0.023)	0.001 (0.035)	0.002 (0.031)
Terrorist Attacks after 10 years old	-0.138*** (0.047)	-0.115** (0.052)	-0.072 (0.062)	-0.048 (0.060)
Mean of Dependent Variable	8.811	8.811	3.355	3.355
Observations	856	856	856	856
Controls	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 5 years old, and between 6 and 9 years old, separately, on Socioemotional Skills Scores. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table 3.7: Effects of Terrorist Attacks on Alcohol Consumption by Childhood periods

	Ever tried alcohol		Age of first drink	
	(1)	(2)	(3)	(4)
Terrorist Attacks 0-5 years old	0.013 (0.009)	0.014 (0.009)	-0.040 (0.104)	-0.037 (0.087)
Terrorist Attacks 6-9 years old	0.037*** (0.013)	0.040*** (0.014)	-0.031 (0.087)	-0.025 (0.085)
Terrorist Attacks after 10 years old	0.054* (0.029)	0.063** (0.030)	0.159 (0.195)	0.159 (0.198)
Mean of Dependent Variable	0.643	0.643	12.84	12.84
Observations	870	870	495	495
Controls	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 5 years old, and between 6 and 9 years old, separately, on Alcohol Consumption Outcomes. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table 3.7 presents mixed evidence regarding the impact of early childhood shocks. Column 2 provides weak evidence suggesting that early-on shocks do not appear to increase the probability of drinking, while violence exposure during childhood appears to increase it. This result is surprising given the previous evidence on how early childhood environment affects future outcomes. However, it also highlights that exposure to violence

matters at the time when the children are first deciding whether to engage in risky health behaviors. Therefore, this could explain why exposure matters when it happens later in life. However, in Column 4, the results for the age of the first drink are not statistically. Due to the limited sample size for this outcome, I lack the statistical power to test if the estimators differ significantly.

Finally, to address the concerns about the assumption related to the timing of the exposure, I estimate the effects using different age cutoffs instead of 5. Figure C.6 shows that most coefficients are significant, indicating that the total difficulties score and the probability of drinking increase, mainly when the early childhood years are included. Notably, the figure suggests a lack of impact when exposure occurs for only a few years, consistent with the results on the duration of exposure discussed earlier.

3.5.4 Robustness checks and Alternative Specifications

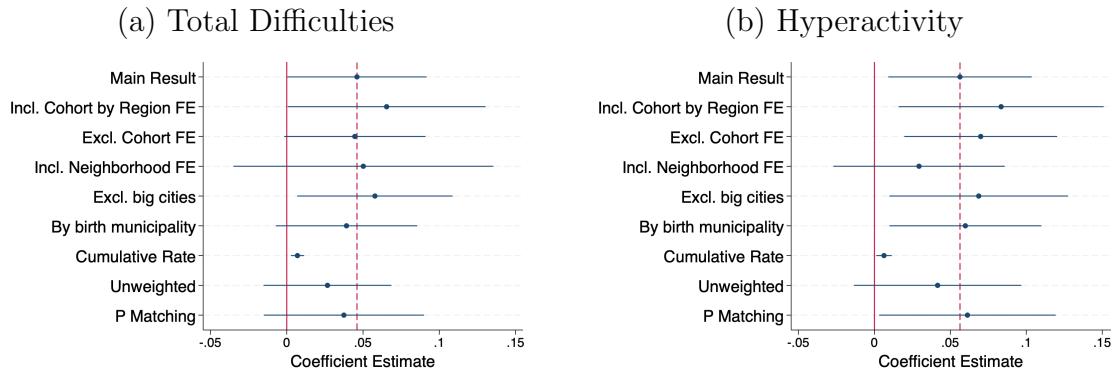
This section provides a brief summary of a number of robustness checks and alternative specifications estimated to address several concerns about the main specification.

Robustness

The initial three specification checks test whether the results are robust to different ways of controlling for differences across cohorts. The first check includes a cohort-by-region fixed effect that controls for region-specific shocks instead of accounting for national shocks. It is important to note that including the cohort-by-municipality fixed effect, which is the ideal, is not plausible, as the variation in terrorist attack levels comes from the within-municipality cross-cohort variation. Furthermore, in this specification, the sample size decreases as the children living in Bogota are excluded because it is both a municipality and a region. The second check excludes the cohort-fixed effect and controls

for age only. Finally, the last specification check includes a neighborhood fixed effect that accounts for shocks experienced by adolescents living in the same neighborhood that might influence their outcomes. Figures 3.7 and 3.8 present the results of these robustness checks.²³ Following the main result (i.e., the main specification), the next three estimates demonstrate the insensitivity to the choice of fixed effects.²⁴

Figure 3.7: Robustness Checks on Socioemotional Skills

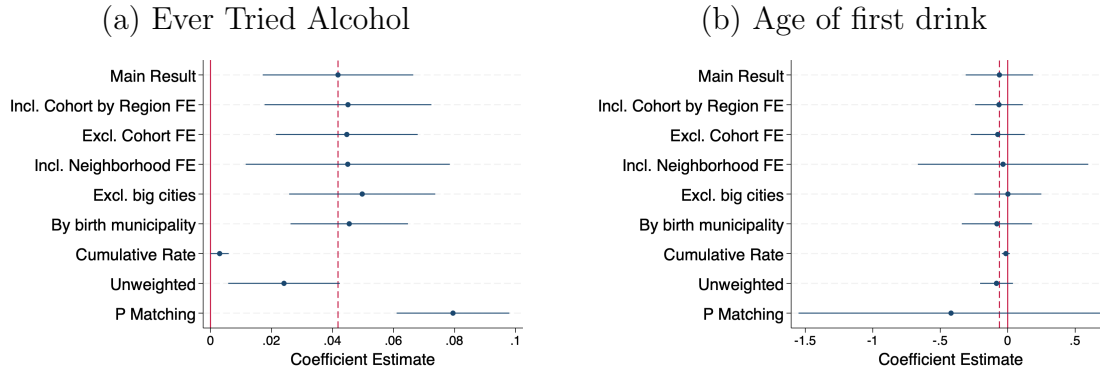


Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Socioemotional Skills Scores for different specifications. Results are weighted by children’s ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level. The “Main Result” estimate (at the top) uses the specification from the even columns of Table 3.4. All other estimates are variations on the baseline model. Estimates 2-4 vary the set of fixed effects. Estimate 5 restricts the sample to small municipalities only, and estimate 6 assigns the rate of terrorist attacks to the birth municipality instead of the residency municipality. Estimate 7 uses the cumulative rate of terrorist attacks between 0 and 9, instead of the average rate, while estimate 8 indicates no weights are used in the regression. Finally, estimate 9 uses a probability matching estimation.

²³Online Appendix Table C.8 presents the results for the other components.

²⁴Due to the survey design, some municipalities contain only a single neighborhood. In these cases, neighborhood and municipality fixed effects are identical. Therefore, I restrict the neighborhood fixed effects analysis to municipalities with multiple neighborhoods, which may explain the lack of significant results, as the sample size decreases.

Figure 3.8: Robustness Checks on Alcohol Consumption



Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Alcohol Consumption Outcomes for different specifications. Results are weighted by children’s ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level. The “Main Result” estimate (at the top) uses the specification from the even columns of Table 3.5. All other estimates are variations on the baseline model. Estimates 2-4 vary the set of fixed effects and use of the instrumental variables strategy. Estimate 5 restricts the sample to small municipalities only, and estimate 6 assigns the rate of terrorist attacks to the birth municipality instead of the residency municipality. Estimate 7 uses the cumulative rate of terrorist attacks between 0 and 9, instead of the average rate, while estimate 8 indicates no weights are used in the regression. Finally, estimate 9 uses a probability matching estimation.

The fourth specification check addresses the concern of whether terrorist attacks in the municipality should impact children in big metropolitan areas. In these sizable cities, a terrorist attack might predominantly affect the children living in the area where it took place, rather than affecting all children throughout the city. To address this concern, I restrict the sample to the municipalities with less than 1 million residents, which excludes the country’s four largest cities (Bogotá, Medellín, Cali, and Barranquilla). However, doing this reduces the sample size further, potentially impacting the significance of the results and diminishing the estimates’ statistical power. The fifth estimate in each panel

in Figures 3.7 and 3.8 shows that the estimates are not sensitive to excluding big cities.

Fifth, within the sample, 13.9% of the children migrated from their birthplace to another location. This notable migration percentage highlights the need to consider migration, which could affect the results. Notably, in Colombia, individuals often migrate from more rural and high-intensity conflict areas to urban and less intense conflict areas. Consequently, the average rate of terrorist attacks is recalculated using the birthplace rather than the municipality of residence in 2010. The main specification outlined in 7 is then re-estimated using this new rate of terrorist attacks. The sixth estimate in each panel in Figures 3.7 and 3.8 shows that the estimates are not sensitive to accounting for migration patterns using the birthplace.

Finally, the results are robust when using the cumulative rate of terrorist attacks and changing the weights. The seventh estimate ("Cumulative Rate") in each panel in Figures 3.7 and 3.8 re-estimates the main specification outlined in 1 using the cumulative rate of terrorist attacks experienced by adolescents during childhood. It shows bigger effects when using the cumulative rate instead of the average rate, but the direction is the same. "Unweighted" indicates no weights are used in the regression using the main specification outlined in Equation 7. The results are slightly different, but given that the survey is representative of the country, weights should be accounted for as the main specification does.

Alternative Specifications

There are two main concerns about the main specification. First, whether it drives the results, which will be addressed with the robustness checks, and some evidence against it was provided in the homicides analysis presented earlier. Second, given the difference-in-differences design, whether the appropriate control group was chosen is even more critical for this paper for two reasons: (i) the choice of fixed effects and (ii) the

strong assumption on how the variation in the timing of the exposure to the terrorist attack is used. In particular, not having a good control group, I cannot fully capture the unobservable factors, such as a municipality being more likely to experience violence or differences across cohorts.

Therefore, in addition to the nearest neighbor match discussed above, another alternative specification is estimated to test the results' sensitivity to the control group's choice. Following (Jacobson and Royer, 2011), the predicted propensity for terrorist attacks is calculated using time-invariant characteristics of the municipalities. Then, the sample is restricted to the municipalities with propensities of violence between 0.1 and 0.9. This approach guarantees that the observations left correspond to more comparable places, at least on their violence levels. Using this new sample, the main specification outlined in 1 is re-estimated.

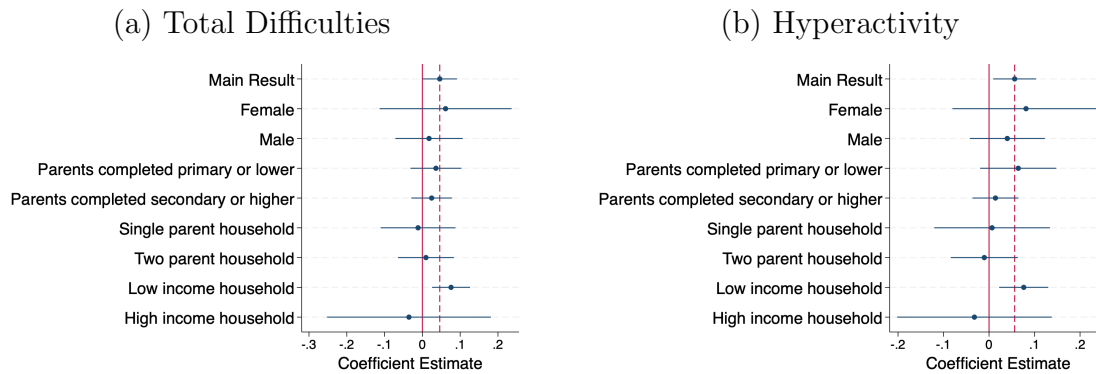
The last estimates in Figures 3.7 and 3.8 represent the specification checks. "P Matching" indicates that the sample is limited to the municipalities with propensities of violence between 0.1 and 0.9, as explained before. The results suggest the same direction of the effects, but some of the point estimates are smaller. Also, the confidence intervals are wide, which is related to the fact that this analysis is restricted to municipalities with propensities of violence between 0.1 and 0.9, which reduces the sample size considerably. However, overall, the results are consistent with those found using the preferred specification.

3.5.5 Heterogeneous results

Having shown that exposure to violence during childhood impacts socioemotional skills and risky health behaviors, this section explores heterogeneity in these estimates across gender, caregivers' education level, household composition, and household income.

Figure 3.9 shows that the detrimental impacts on socioemotional skills do not differ across any of the variables analyzed. Notably, it appears that males, individuals with more educated caregivers, and those living in low-income households are the most affected. However, the differences between the coefficients are not statistically significant; hence, these findings are suggestive evidence.²⁵

Figure 3.9: Heterogeneous Impacts on Socioemotional Skills



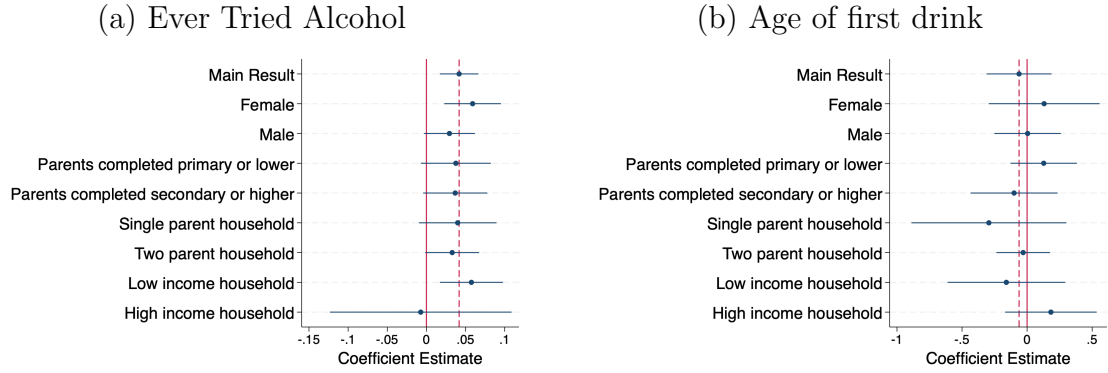
Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Socioemotional Skills Scores separated based on characteristics of the children, their parents, and their household. Results are weighted by children's ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Figure 3.10 shows that the probability of drinking alcohol (Panel A) increases for females and those living in low-income households with exposure to terrorist attacks. However, the differences between these coefficients and those of males and high-income households are statistically insignificant. Even though these findings are suggestive evidence, they provide an interesting pattern. Specifically, men drink more (67.6%) than women (60%), but the exposure to conflict appears to trigger more alcohol consumption.

²⁵Online Appendix Figure C.9 presents the results for the other components.

Moreover, exposure prompts consumption at even earlier ages (Panel B), which has even worse long-term consequences due to its persistence and effects on health.

Figure 3.10: Heterogeneous Impacts on Alcohol Consumption



Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Alcohol Consumption Outcomes separated based on characteristics of the children, their parents, and their household. Results are weighted by children's ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

3.5.6 Placebo test

To verify the validity of the empirical strategy, I check whether the average rate of terrorist attacks between 2 and 7 years before the children's birth year affects the outcomes. The year immediately prior is excluded because it is the in-utero period, and the evidence has shown that in-utero exposure to violence has negative consequences on the children's development (Camacho, 2008). Therefore, the following equation is estimated:

$$y_{imt} = \beta_0 + \beta_1 V_{mt}^{2-7 \text{ before}} + \mathbf{X}_{imt}\beta_2 + \mathbf{Z}_{imt}\beta_3 + \mathbf{C}_{imt}\beta_4 + \mu_m + \delta_t + \varepsilon_{imt} \quad (8)$$

The coefficient of interest is also β_1 , which captures the effect of the terrorist attacks before the teenager was born. Tables 3.8 and 3.9 show the results of the placebo test on socioemotional skills and alcohol-related outcomes. The point estimates are smaller than the main effects in Tables 3.4 and 3.5, and not statistically different from zero. Overall, there are no significant effects, which indicates that terrorist attacks occurring before birth do not affect the outcomes of the adolescent.²⁶

Table 3.8: Placebo Effects of Terrorist Attacks on Socioemotional Skills

	Total Score	Hyperactivity
	(1)	(2)
Terrorist Attacks 2-7 years before	-0.030 (0.019)	-0.020 (0.024)
Mean of Dependent Variable	8.811	3.355
Observations	856	856
Controls	Yes	Yes
Cohort FE	Yes	Yes
Municipality FE	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of terrorist attacks between 2 and 7 years the child was born on Socioemotional Skills Scores. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

²⁶Online Appendix Table C.7 presents the results for the other components.

Table 3.9: Placebo Effects of Terrorist Attacks on Alcohol Consumption

	Ever tried alcohol	Age of first drink
	(1)	(2)
Terrorist Attacks 2-7 years before	-0.015 (0.012)	-0.069 (0.064)
Mean of Dependent Variable	0.643	12.84
Observations	870	495
Controls	Yes	Yes
Cohort FE	Yes	Yes
Municipality FE	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of terrorist attacks between 2 and 7 years the child was born on Alcohol Consumption Outcomes. All models include controls for children’s biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children’s ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

3.6 Conclusion

This paper extends the literature on the impact of exposure to conflict during childhood on medium outcomes by investigating previously unstudied medium-term effects on socioemotional skills and risky health behaviors during adolescence, which are a new set of outcomes. The findings suggest that exposure to violence has significant and lasting consequences for children’s socioemotional skills and risky health behavior. Notably, the severity and duration of the exposure are key to understanding these adverse effects. The results also underscore that exposure earlier in life has worse consequences and affects even

more medium-term outcomes. Overall, these results highlight the importance of childhood and how exposure for an extended period and high intensity might have long-term effects.

While the paper has limitations regarding data quality and endogeneity concerns, the results provide valuable insights into the impact of terrorist attacks in countries like Colombia, where violence has been pervasive for many years. Since 2000, over 43 countries around the globe have experienced civil war and violence (Marshall, 2020), and understanding how to mitigate the negative effects of those conflicts is of first-order importance. Public discussions center on the economic consequences of the armed conflict, but less is known about socioemotional skills and risky health behaviors during adolescence. In particular, the results indicate that exposure to violence can alter children's life trajectories, with negative implications for their health, education, or employment outcomes.

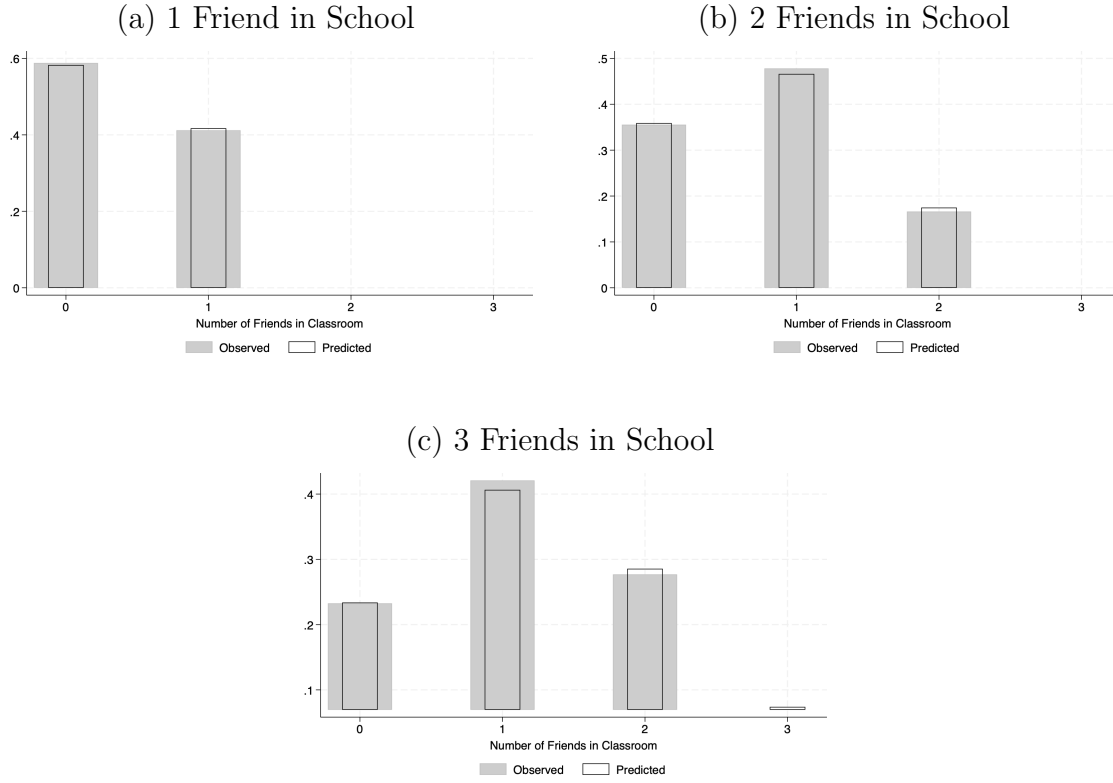
Finally, given the literature on the impacts of socioemotional skills and risky health behaviors on adulthood, these results raise essential policy concerns about properly addressing these adverse effects. Future research should explore the underlying mechanisms through which exposure to violence affects children and identify ways to support affected children and their families through public policy interventions. Additionally, it should explore how the effects of conflict mediate the relationship between non-cognitive skills and risky health behaviors.

Appendix A

Appendix for Chapter 1

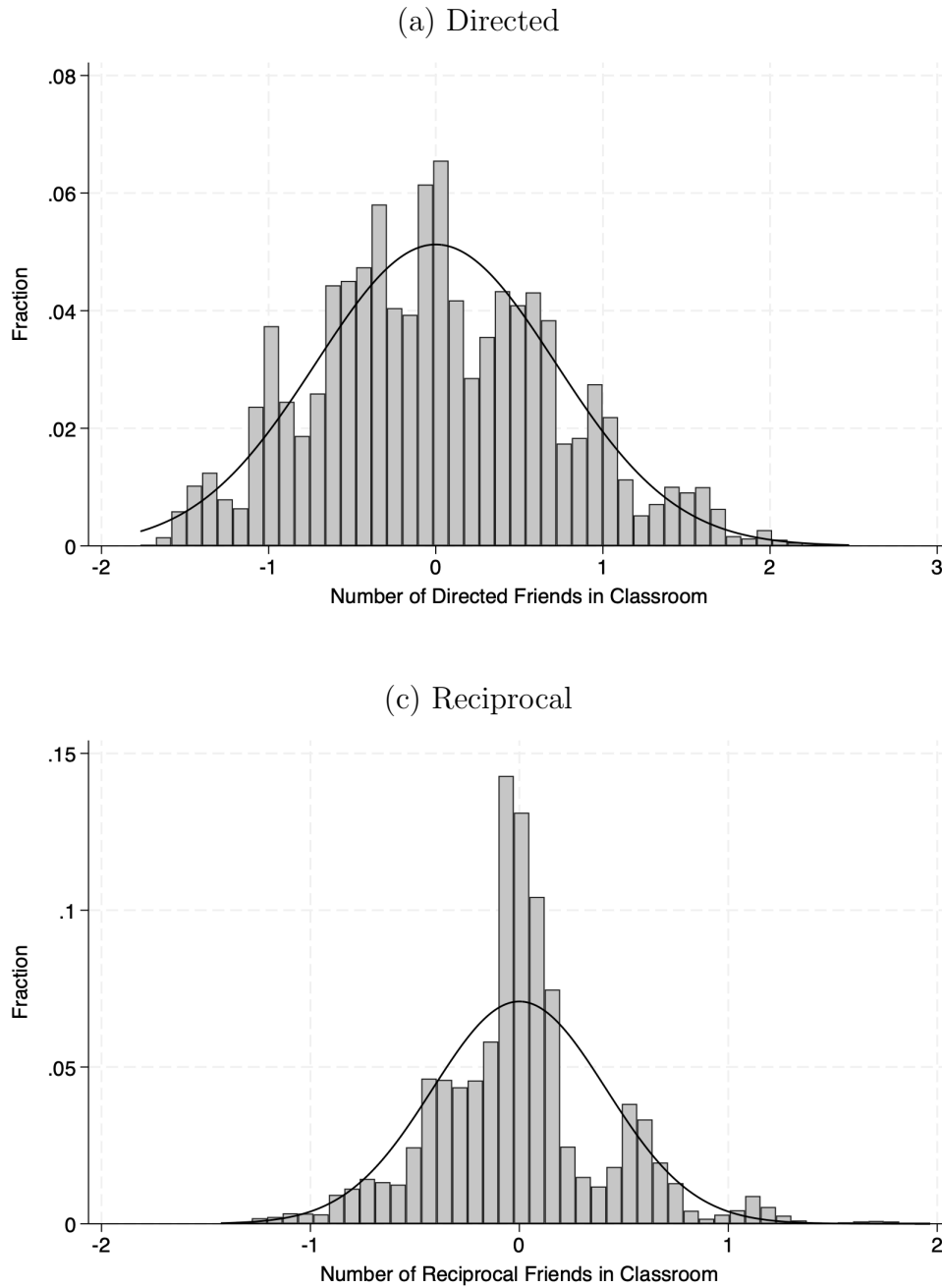
A.1 Supplemental Figures and Tables

Figure A.1: Observed Vs Expected Probability of Friends in Classroom



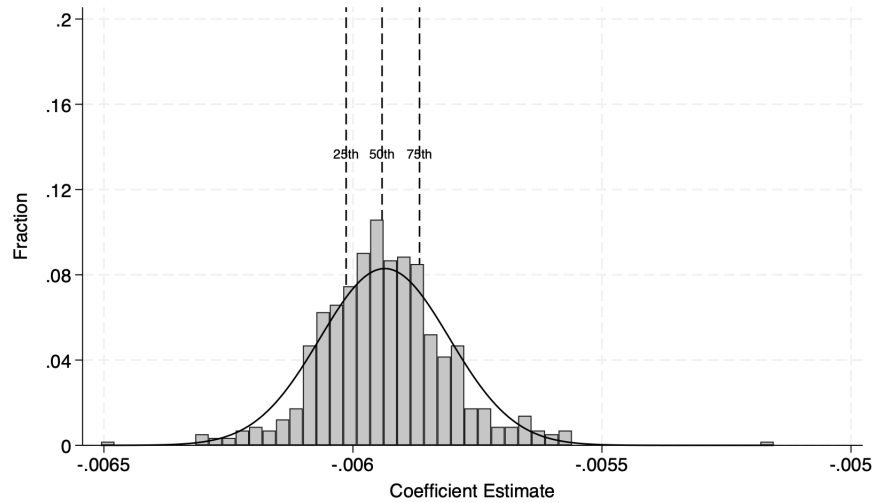
Notes: The figure above shows the observed and predicted probability of being randomly assigned to a classromm with k friends in the classroom, conditional on the number of friends in the school. The expected probabilities were calculated using the hypergeometric distribution with the school- and student-specific parameters.

Figure A.2: Residual Number of Friends Across School-Grades



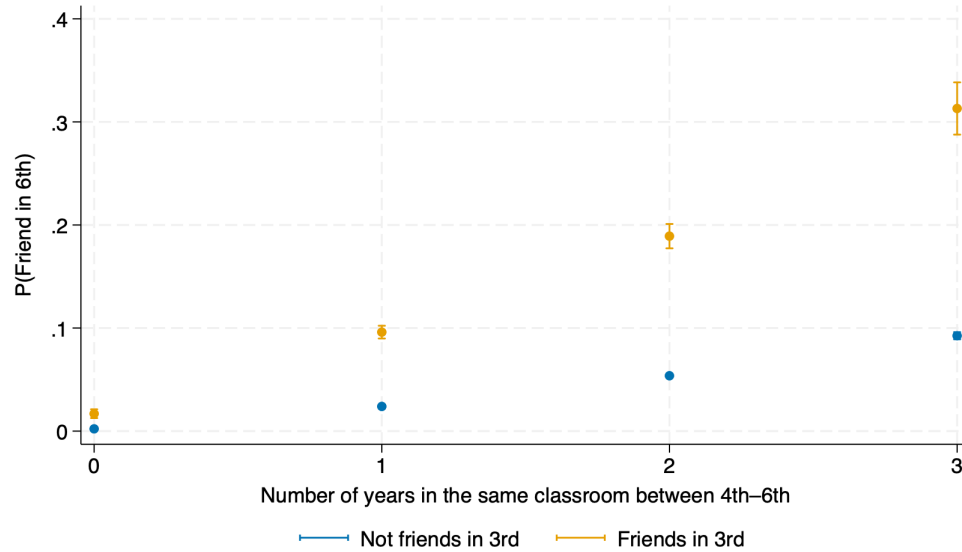
Notes: The figure above represents the residualized distribution of the number of friends in the classroom across school-grades, conditional on school-by-grade fixed effects. All regressions are limited to schools with at least two classrooms per grade. The overlaid curve represents the normal distribution.

Figure A.3: Distribution of Coefficients of the Effect of Friends in the Classroom on Math Scores



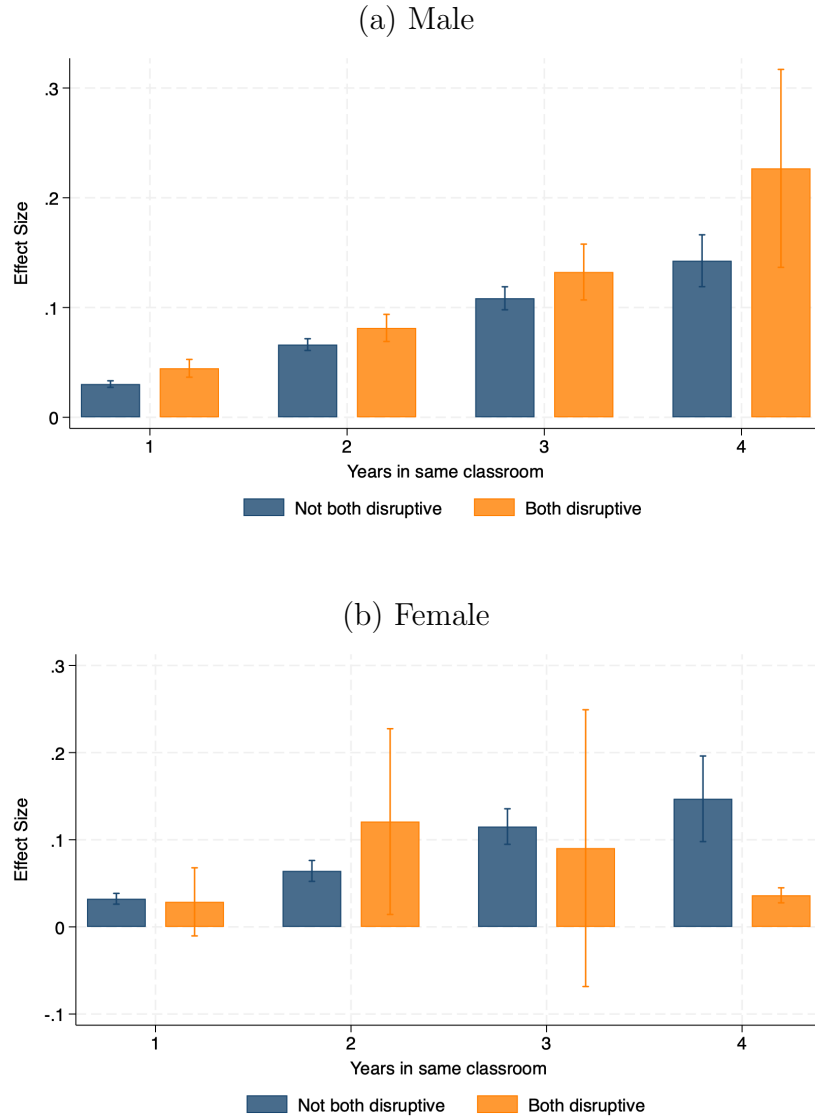
Notes: The figure above represents the distribution of the coefficients of leave-one-out regressions using the main specification. Each line represents the main specification with one school-grade committed from the sample. All regressions are limited to schools with at least two classrooms per grade. 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: Impacts on the Probability of Friendship Persistence by Years of Exposure



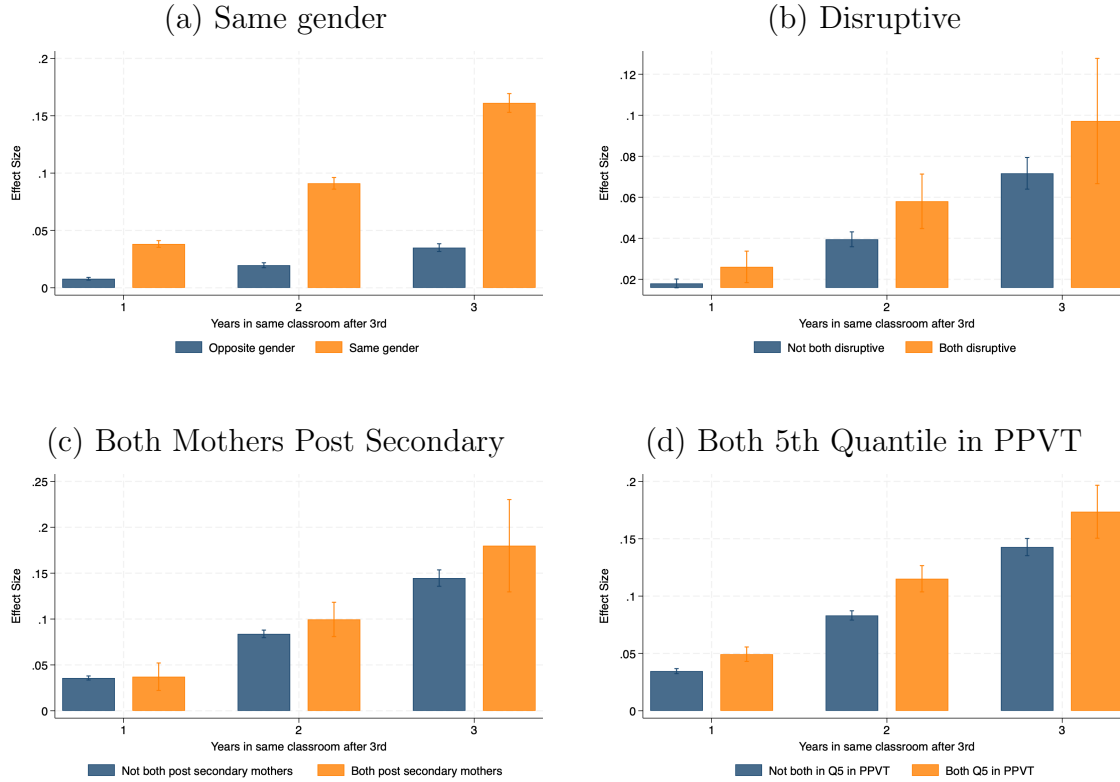
Notes: The figures above report non-linear estimates from regressions of binary indicators for the number of years of exposure between 4th and 6th grade. It shows the results for friendship persistence only for directed links, separately for those who were friends in third grade and those who were not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. 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 A.5: Impacts on Friendship Formation by Years of Exposure and Homophily (Disruptive)



Notes: The figures above report estimates of the predicted increase of the probability that students i and j have a directed link. The estimates were obtained from regressions of binary indicators of the number of years students i and j are randomly assigned to the same classroom, interacted with a binary indicator of homophily: pairs of students that are both disruptive compared with pairs where one is disruptive and one is not, separately by gender. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. All regressions are limited to schools with at least two classrooms per grade. Vertical lines represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Figure A.6: Impacts on Friendship Persistence by Years of Exposure and Homophily



Notes: The figures above report estimates of the predicted increase of the probability that students i and j directed link persists. The estimates were obtained from regressions of binary indicators of the number of years students i and j are randomly assigned to the same classroom interacted with a binary indicator of homophily among four characteristics: (i) same-gender pairs compared with opposite-gender pairs; (ii) pairs of students that are both disruptive compared with pairs where one is disruptive and one is not; (iii) pairs of students that both have post-secondary educated mothers compared with pairs where one does and one does not; and (iv) pairs of students where both are in the highest quintile of the PPVT distribution compared with pairs where one is and one is not. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. All regressions are limited to schools with at least two classrooms per grade. Vertical lines represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the school level.

Table A.1: Impacts on Persistence of Reciprocal Friendships

	Not Friends in 3rd		Friends in 3rd	
	(1)	(2)	(3)	(4)
Same classroom in 4th	0.001*** (0.000)		0.013 (0.008)	
Same classroom in 5th	0.003*** (0.000)		0.022*** (0.007)	
Same classroom in 6th	0.012*** (0.001)		0.110*** (0.009)	
Years in same classroom after 3rd		0.005*** (0.000)		0.049*** (0.005)
Mean of Dependent Variable	0.005	0.005	0.053	0.053
Observations	504661	504661	4456	4456
School FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions on a binary indicator of whether students i and j have a reciprocal friendship in sixth grade, separately for pairs of students who were friends in third grade and pairs that were not. A reciprocal friendship is defined as one in which both student i and student j nominate each other. Columns (1) and (3) present the results using binary indicators for whether the students were randomly assigned to the same classroom in each grade between 4th and 6th grade. Columns (2) and (4) present the results using the total number of years the students were randomly assigned to the same classroom between 4th and 6th grade. All regressions are limited to schools with at least two classrooms per grade. *** 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 A.2: Impacts on Friendships on Other Non-Cognitive Skills

	Directed	Reciprocal	Followers	Rejecters
	(1)	(2)	(3)	(4)
A. Growth Mindset				
No. of Friends in Classroom	-0.015 (0.017)	0.003 (0.026)	0.017 (0.019)	-0.022 (0.022)
Mean of Dependent Variable	-0.001	-0.001	-0.001	-0.001
Treatment Effect of 1SD increase	-0.012	0.002	0.016	-0.016
Observations	7535	7535	7535	7535
B. Grit				
No. of Friends in Classroom	-0.012 (0.015)	-0.029 (0.027)	0.017 (0.019)	-0.002 (0.018)
Mean of Dependent Variable	-0.002	-0.002	-0.002	-0.002
Treatment Effect of 1SD increase	-0.009	-0.016	0.016	-0.001
Observations	7535	7535	7535	7535
Controls	Yes	Yes	Yes	Yes
School-by-grade FE	Yes	Yes	Yes	Yes

Notes: The table reports estimates from regressions of the number of friends in the classroom on non-cognitive skills, conditional on the school-by-grade fixed effects and the number of friends in the school. Each column represents one of the friendship types analyzed: directed links (i.e., student i nominates student j , regardless of whether student j reciprocates), reciprocal friendships (i.e., both student i and student j nominate each other), followers (i.e., those who nominated a student but were not reciprocally nominated) and rejecters (i.e., those who are nominated by a classmate but do not reciprocate). All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, and age squared. *** 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-by-grade level.

A.2 Homophily Descriptives

In this appendix, we estimate additional descriptive analysis of homophily. First, Appendix Table A.3 examines homophily in directed links using both observed and expected link counts. Column (1) reports the average number of directed links. Column (2) reports the average share of each student's friends who share the same characteristic as the student. For example, among female students, 90.40% of their friends are also female, indicating substantial within-gender linking; by contrast, the analogous pattern for maternal education is present but weaker. However, these patterns should not be interpreted as evidence of homophilous preferences because they can arise mechanically from the underlying composition of the population. For instance, if all students had mothers with post-secondary education, then 100% of the links would be within that group. To separate composition from preference, Column (3) reports population shares for each characteristic in the sample. Comparing Columns (2) and (3), the share of same-type friends exceeds the corresponding population share, with the largest gaps observed for gender. Under random sorting, the shares in Column (2)) would match the population share in Column (3). Their systematic excess indicates homophily in friendship formation.

Table A.3: Directed links by gender and by maternal education (3rd grade)

	No. of Friends	Share same group	Group share
A. Child's Gender			
Female	2.70	90.40	48.30
Male	2.68	87.75	51.70
B. Maternal Education			
Female			
Some Primary	2.74	45.55	40.16
Some Secondary	2.73	55.32	50.95
Post Secondary	2.78	19.97	8.89
Male			
Some Primary	2.72	47.19	38.44
Some Secondary	2.71	57.60	52.03
Post Secondary	2.73	14.87	9.53

Notes: The Table reports statistics about homophily in directed links. A directed link is defined as one in which student i nominates student j , regardless of whether student j reciprocates. The "No. of Friends" columns is the average number of friends a student has in the school. The "Share same group" column presents the mean of the percentage of directed friends from the same gender or same maternal education as their own. The "Group share" column provides the distribution of students by gender and maternal education in the full sample.

Second, we compare the percentage of directed links in which both students share the

same characteristic in the data to the percentage expected under random assignment. Assuming that all students form the same number of links, we compute the expected share of homophilous links based on the distribution of characteristics in the population.¹ Appendix Table A.4 compares the expected and observed percentages of homophilous friendships. For gender, we would expect approximately 50% of links to be between students of the same gender, but we observe around 97%. For maternal education, the expected share of homophilous links is around 43%, while the observed share is approximately 48%. These results are consistent with earlier descriptive findings and suggest that homophily is more pronounced by gender than by maternal education.

Table A.4: Expected and Actual Homophilous Friendships

	Homophilous Friendships	
	Expected	Observed
A. Child's Gender		
Female	23.45	52.55
Male	26.60	44.71
B. Maternal Education		
Female		
Some Primary	16.00	18.53
Some Secondary	26.02	28.03
Post Secondary	0.81	1.94
Male		
Some Primary	15.00	16.96
Some Secondary	26.88	30.18
Post Secondary	0.89	1.53

Notes: The Table presents the expected and observed percentage of homophilous friendships based on the percentage of people with a given characteristic in the sample.

¹If there are n characteristics, the expected fraction of homophilous friendships is $\sum_{i=1}^n p_i^2$ where p_i is the fraction of students with characteristic i in the sample.

A.3 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 A.5 and Table A.6 show summary statistics of each of the tests applied by grade for math and executive function, respectively.

Table A.5: Summary Statistics for Math Test Score Components

	N	Mean	SD
A. 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
B. 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
C. 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 in 4th grade. All the correlations are significant at the 1 percent level.

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 A.6: Summary Statistics for Executive Function Score Components

	N	Mean	SD
A. 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 in 4th grade. All the correlations are significant at the 1 percent level.

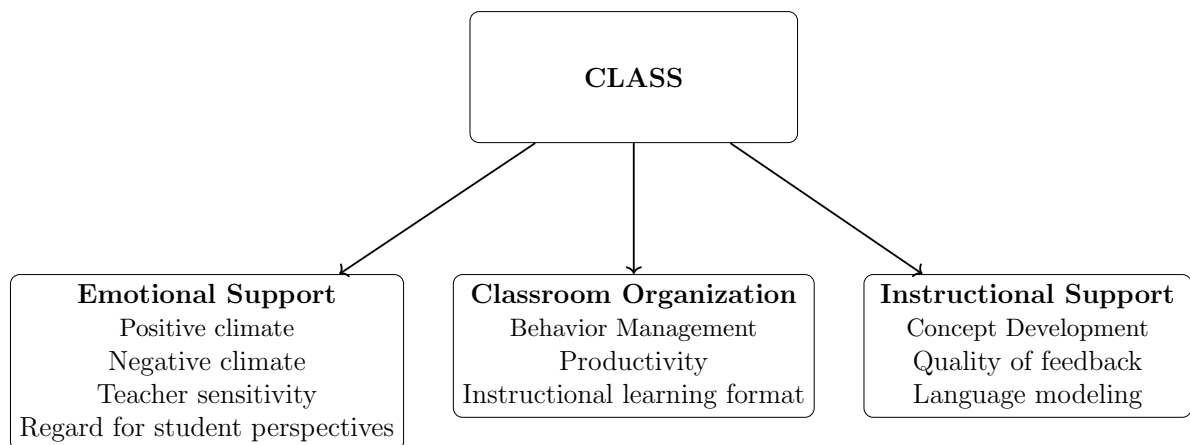
A.4 Application of the CLASS in Ecuador

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 A.7 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.

Figure A.7: 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.

Application in Ecuador

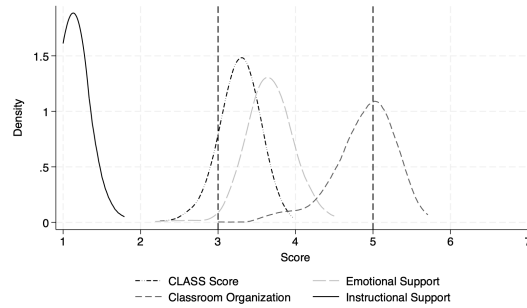
We filmed all teachers for an entire school day (from approximately eight to one in the afternoon for morning schools and from two to six in the afternoon for the afternoon schools). Teachers only knew on what day they would be filmed until the day itself. Following CLASS protocols, we discarded the 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.² Additionally, during the entire process, we interacted extensively with the developers of the CLASS at the University of Virginia.

Figure A.8 graphs univariate densities of the distribution of CLASS score and each domain in 4th 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.

²Based 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

Figure A.8: 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 in 4th grade.

Table A.7 shows the correlation between the total CLASS score and each of its components. It shows that the scores of the components 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 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.

Table A.7: Correlation across CLASS scores

	Total Score	Emotional	Classroom	Instructional
	Total Score	Support	Organization	Support
	(1)	(2)	(3)	(4)
Total Score	1.000			
Emotional Support	0.869	1.000		
Classroom Organization	0.871	0.589	1.000	
Instructional Support	0.622	0.400	0.382	1.000

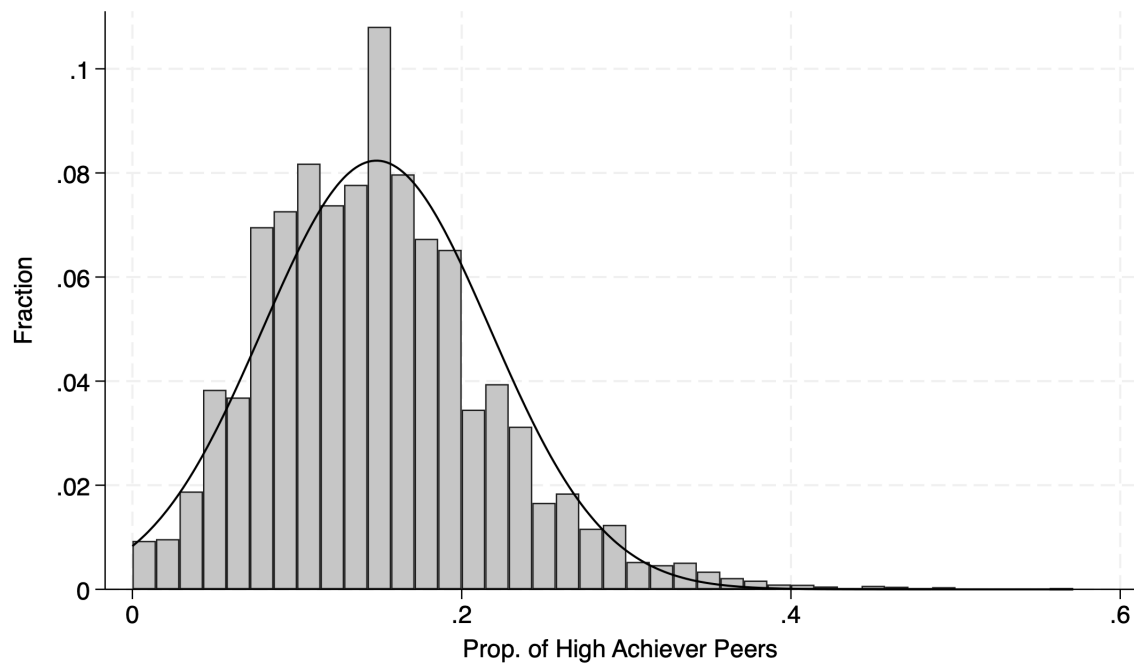
Notes: The table presents the results from pairwise correlations between CLASS score and its components collected in 4th grade. All the correlations are significant at the 1 percent level.

Appendix B

Appendix for Chapter 2

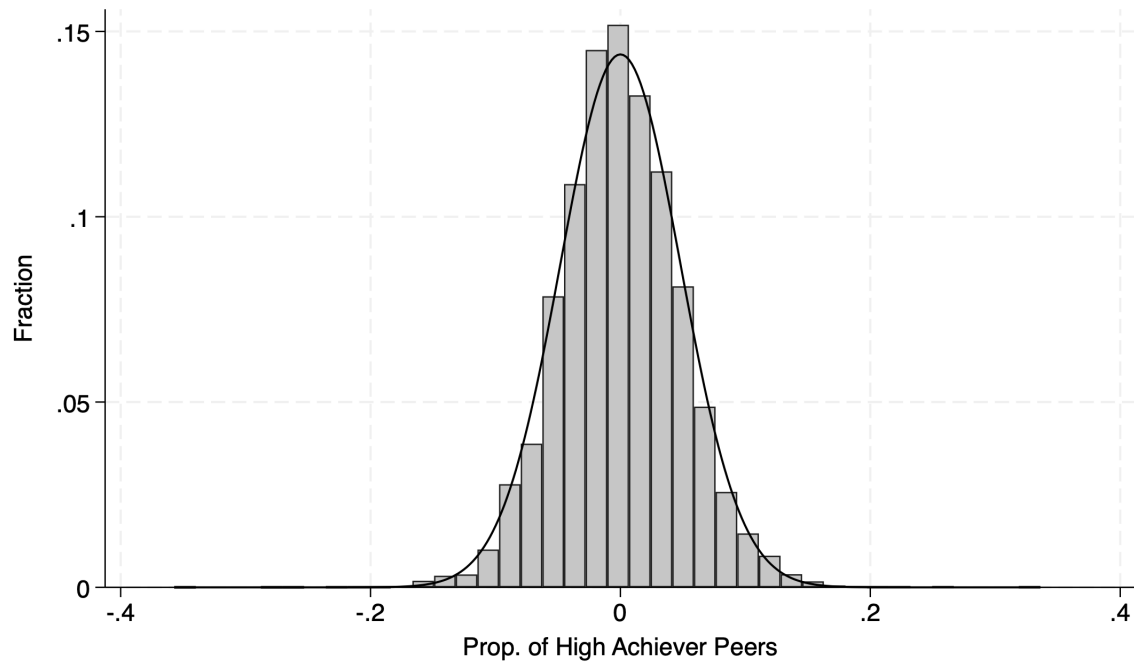
B.1 Supplemental Figures and Tables

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



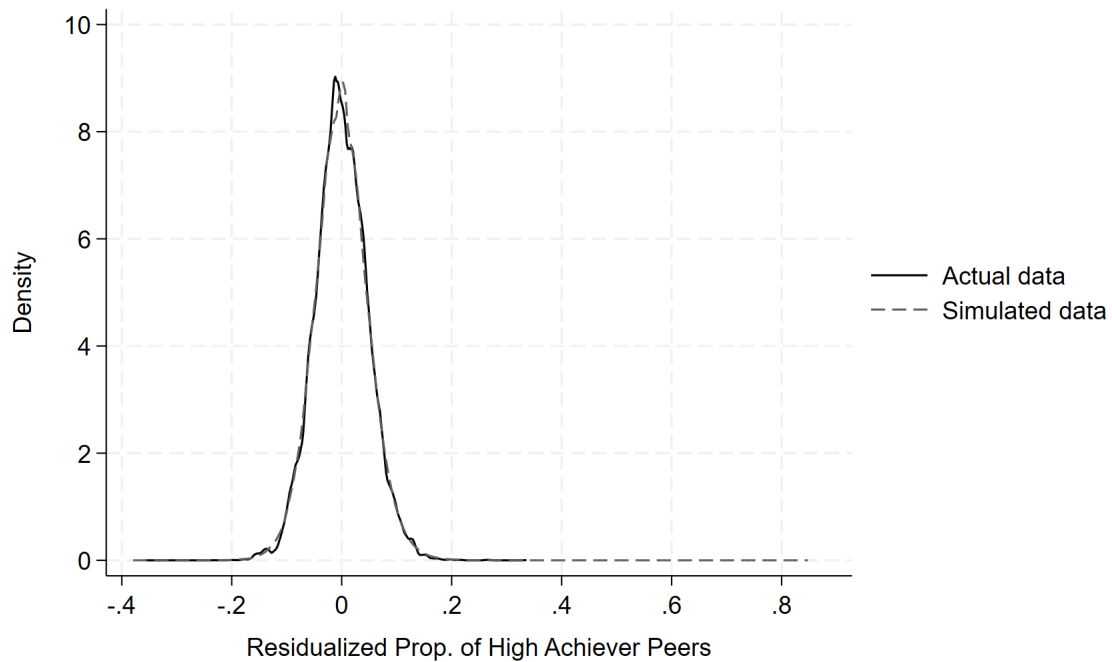
Notes: The figure above high achiever peer share distribution across school-grades. The overlaid curve represents the normal distribution.

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



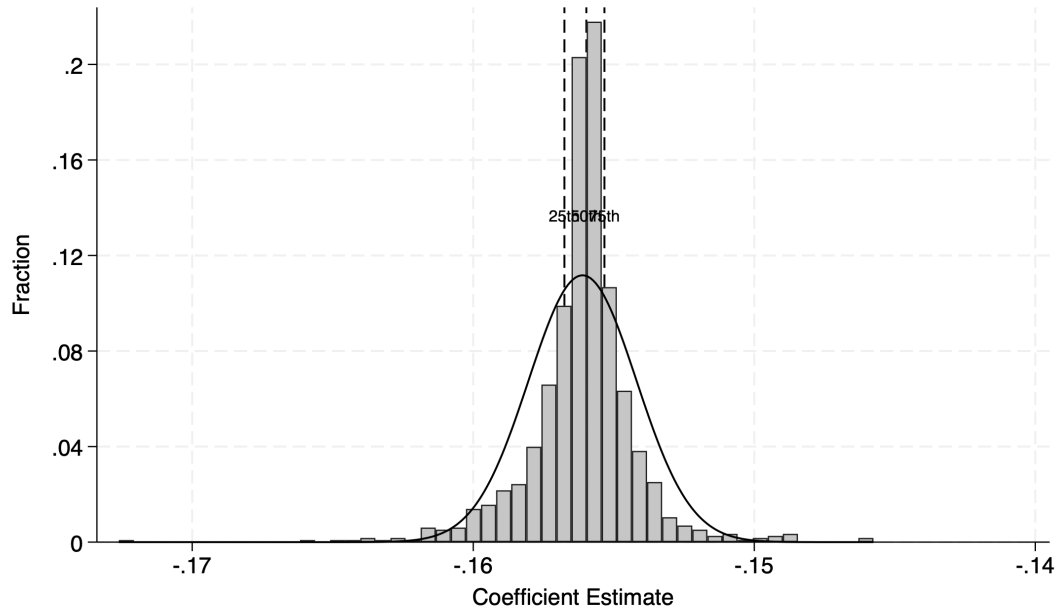
Notes: The figure above represents the residualized high achiever peer share distribution across school-grades, conditional on school-by-grade fixed effects. All regressions are limited to schools with at least two classrooms per grade. The overlaid curve represents the normal distribution.

Figure B.3: Simulated and Actual Residual Share of High Achievers Peers Across School-Grades



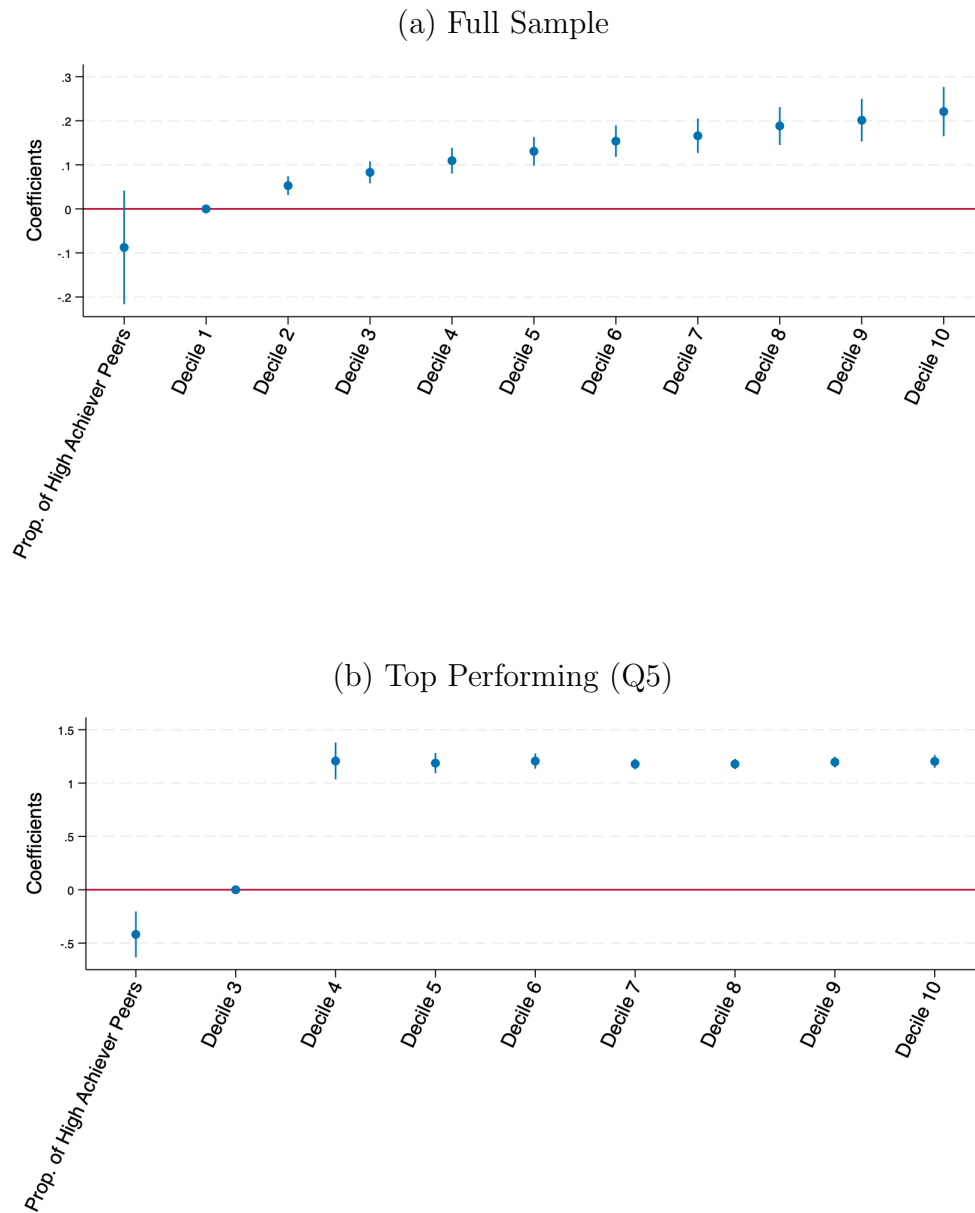
Notes: The figure above represents the residualized high achiever peer share distribution across school-grades, conditional on school-by-grade fixed effects. All regressions are limited to schools with at least two classrooms per grade. The overlaid curve represents the normal distribution.

Figure B.4: Distribution of Coefficients of the Effect of High Achievers on Math Scores



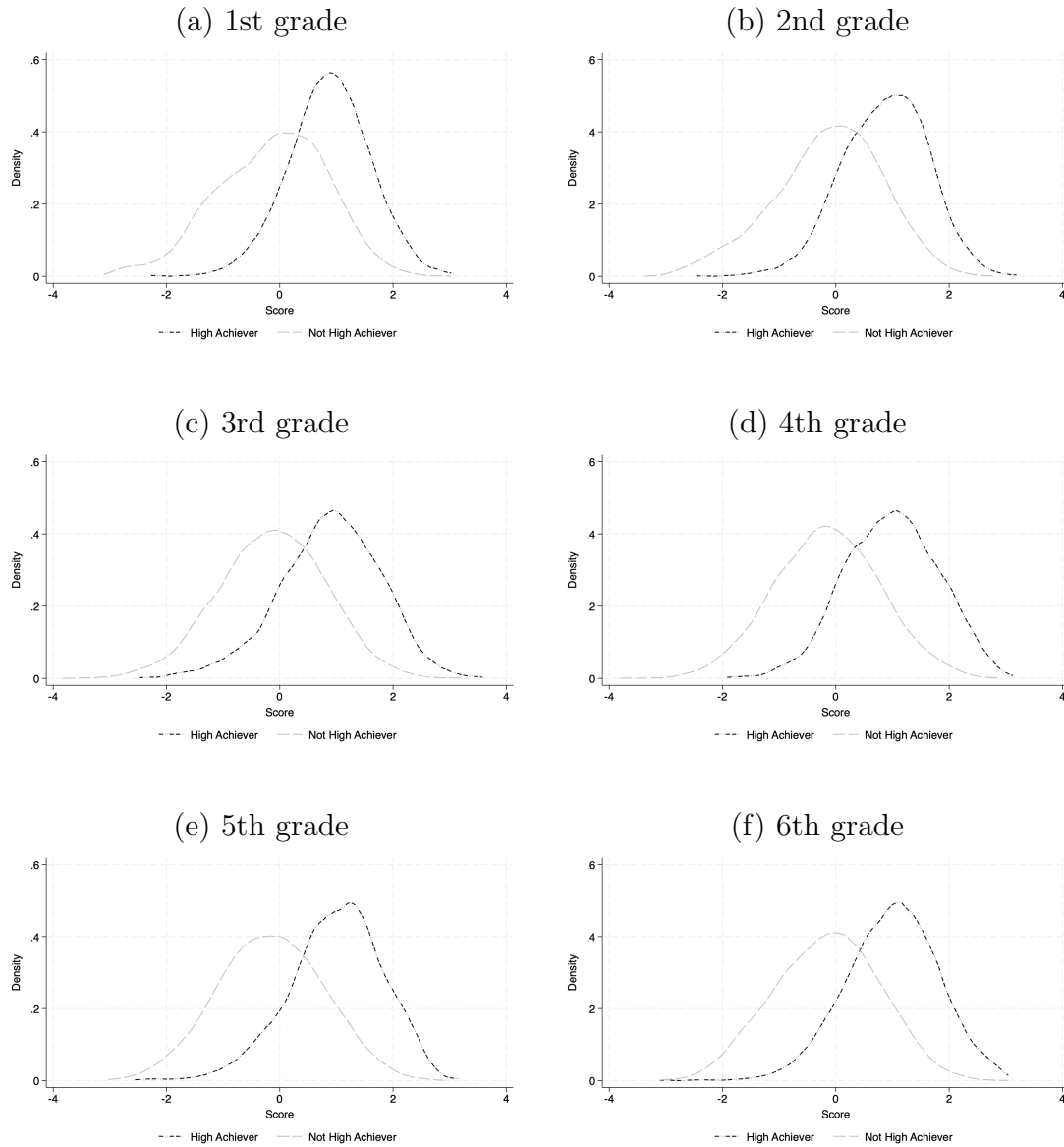
Notes: The figure above represents the distribution of the coefficients of leave-one-out regressions using the main specification. Each line represents the main specification with one school-grade committed from the sample. All regressions are limited to schools with at least two classrooms per grade. 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 B.5: Effects of High Achievers and Classroom Rank Deciles on Math Scores



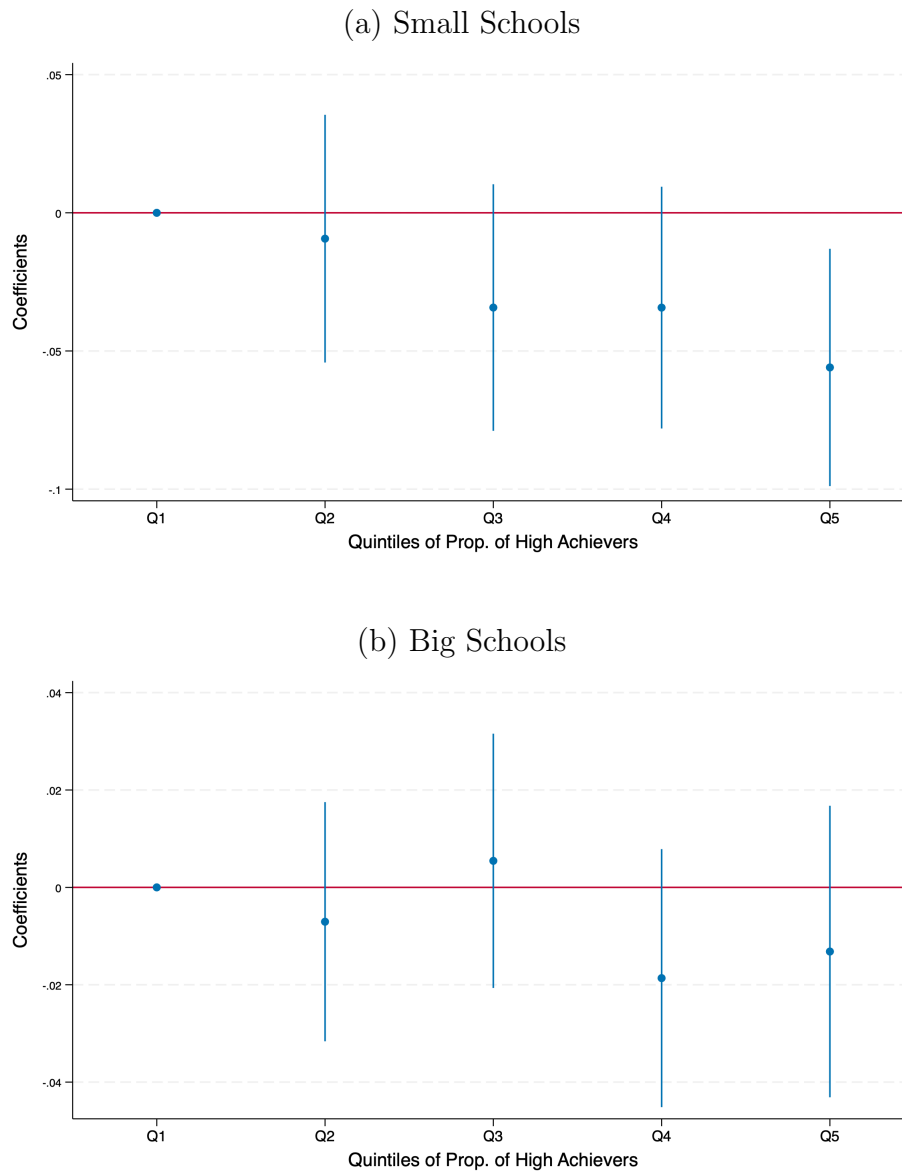
Notes: The figure reports estimates from regressions of the leave-one-out proportion of high-achiever peers and the deciles of classroom rank measure in percentiles (so rank is normalized to be between zero and one) similarly to Carneiro et al. (2025b). All regressions are limited to schools with at 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.

Figure B.6: Distribution of Math Scores of high-achieving students



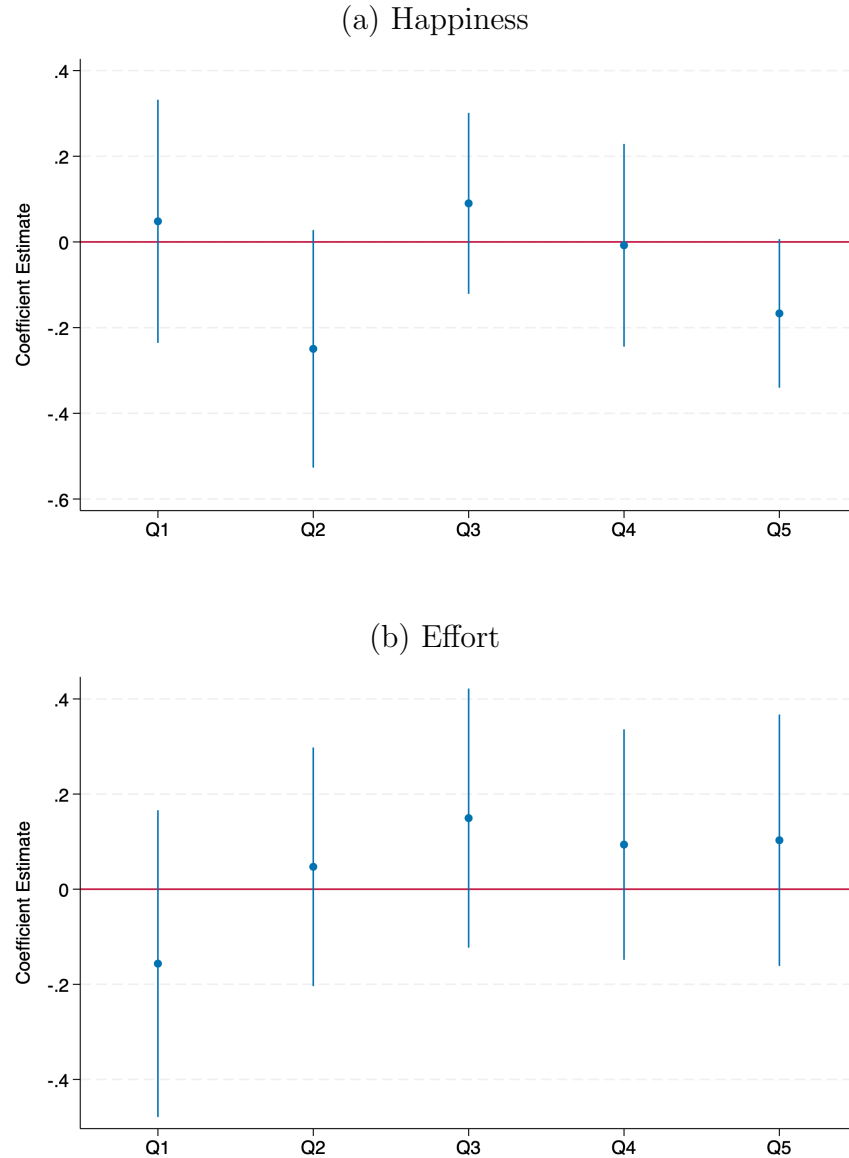
Notes: The figure above presents the univariate densities of the distribution of the math test score, separately for high achievers and non-high achievers based on teacher rankings, by grade.

Figure B.7: Interaction of Non-Linear Effects and School Size



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 B.8: Heterogeneous Effects of High Achievers on Happiness and Effort by Previous Math Test Results



Notes: The figures above report estimates from regressions of the leave-one-out proportion of high-achiever peers on happiness and effort by quintiles of the math test result in the previous year. 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.

Table B.1: Characteristics of Female and Male high-achieving students

	Male	Female	Difference (1)-(2)
	(1)	(2)	(3)
A. Children characteristics			
Age of child (in months) in 2012	61.429 (5.031)	60.938 (4.595)	0.492** (0.222)
Receptive vocabulary score (PPVT)	92.209 (16.673)	91.137 (16.703)	1.071 (0.768)
Proportion who attended preschool	0.640 (0.480)	0.650 (0.477)	-0.010 (0.022)
B. Household Characteristics			
Mother's years of completed schooling	9.822 (3.978)	9.837 (3.910)	-0.015 (0.184)
Father's years of completed schooling	9.480 (4.065)	9.465 (3.968)	0.015 (0.208)
Mother's age	30.798 (6.861)	30.638 (6.360)	0.160 (0.310)
Father's age	35.368 (8.420)	34.314 (7.358)	1.054** (0.412)
Household has piped water in home	0.859 (0.348)	0.826 (0.380)	0.033** (0.017)
Household has flush toilet in home	0.457 (0.498)	0.484 (0.500)	-0.027 (0.023)

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

Table B.2: Variation in Proportion of High Achievers After Removing Fixed Effects

	Mean	SD	Median	Min	Max
A. Raw variables					
Proportion of High Achievers	0.15	0.07	0.14	0.00	0.57
B. Residuals after removing school-by-grade FE					
Residuals of Proportion of High Achievers	-0.00	0.05	-0.00	-0.36	0.35

Notes: The table above reports the raw variation in the leave-one-out proportion of high achievers and the variation that is left after removing school-by-grade fixed effects. The table only includes schools with at least two classrooms per grade.

Table B.3: 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.2 Randomization details

Table B.4: 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.

B.3 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 B.5 and Table B.6 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 B.5: 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 B.6: 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.

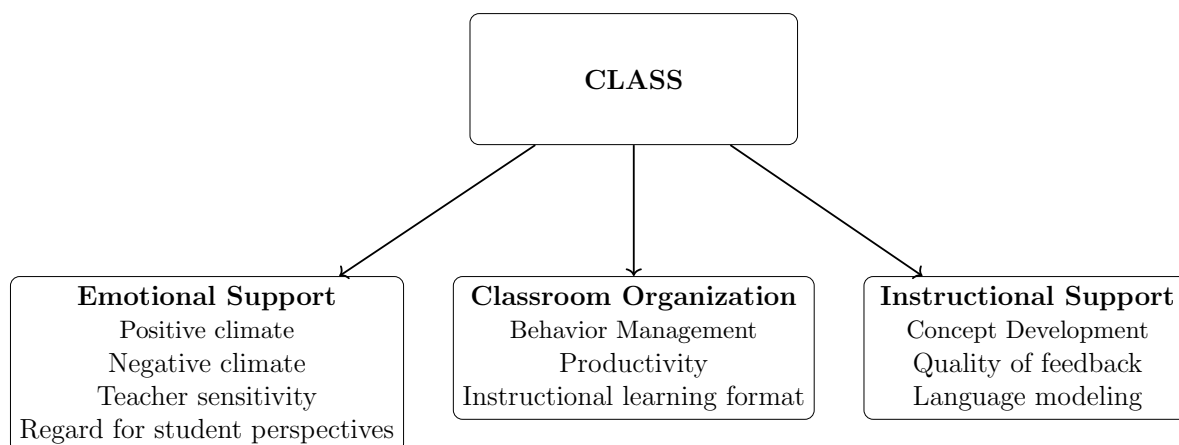
B.4 Application of the CLASS in Ecuador

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 B.9 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.

Figure B.9: 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.

Application in Ecuador

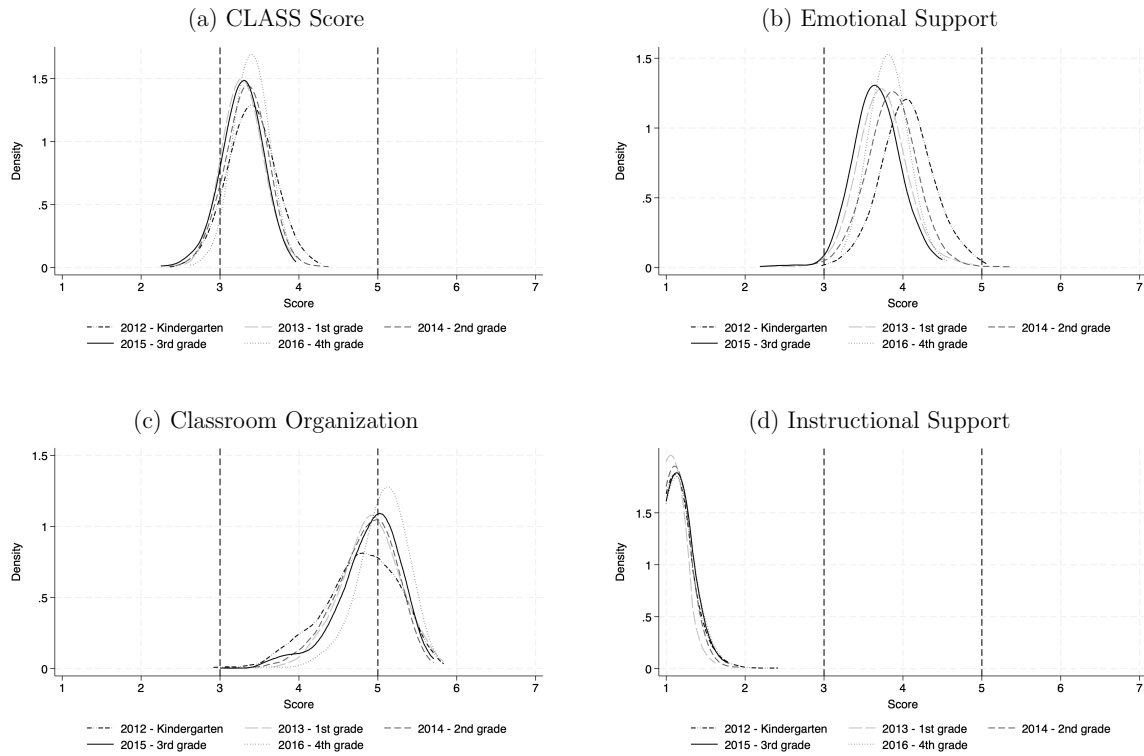
We filmed all teachers for an entire school day (from approximately eight to one in the afternoon for morning schools and from two to six in the afternoon for the afternoon schools). Teachers only knew on what day they would be filmed until the day itself. Following CLASS protocols, we discarded the 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.¹ Additionally, during the entire process, we interacted extensively with the developers of the CLASS at the University of Virginia.

Figure B.10 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.

¹Based 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

Figure B.10: 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 B.7 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 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.

Table B.7: Correlation across CLASS scores

	Total Score	Emotional	Classroom	Instructional
	Total Score	Support	Organization	Support
	(1)	(2)	(3)	(4)
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.

B.5 Extra Robustness Checks

Peer's Parents Education

In this appendix, we estimate the main results in Table 1.5 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 B.8, 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.

Table B.8: 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 (0.058)	-0.106 (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.

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 B.9 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 B.9 shows that around 22% of the children in Ecuador have parents who have finished secondary school.

Table B.9: 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 B.10 shows that the results using the proportion of children that both parents finished secondary school are mostly consistent with the results in Table 1.5 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 B.10: Effects of Peers' Parents Secondary Education on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of Peers with Parents who finished Secondary	-0.048 (0.052)	-0.111* (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 B.11 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 B.12, we restrict our sample to the balanced panel of children and estimate the main equation. We find that the results are similar.

Table B.11: Effects of High Achievers on Attrition

	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 (0.011)	-0.001 (0.031)	0.002 (0.013)	-0.037 (0.041)	0.019 (0.017)	-0.000 (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 B.12: Effects of High Achievers on Cognitive Skills

	Math	Executive Function
	(1)	(2)
Prop. of High Achiever Peers	-0.161** (0.080)	-0.299*** (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

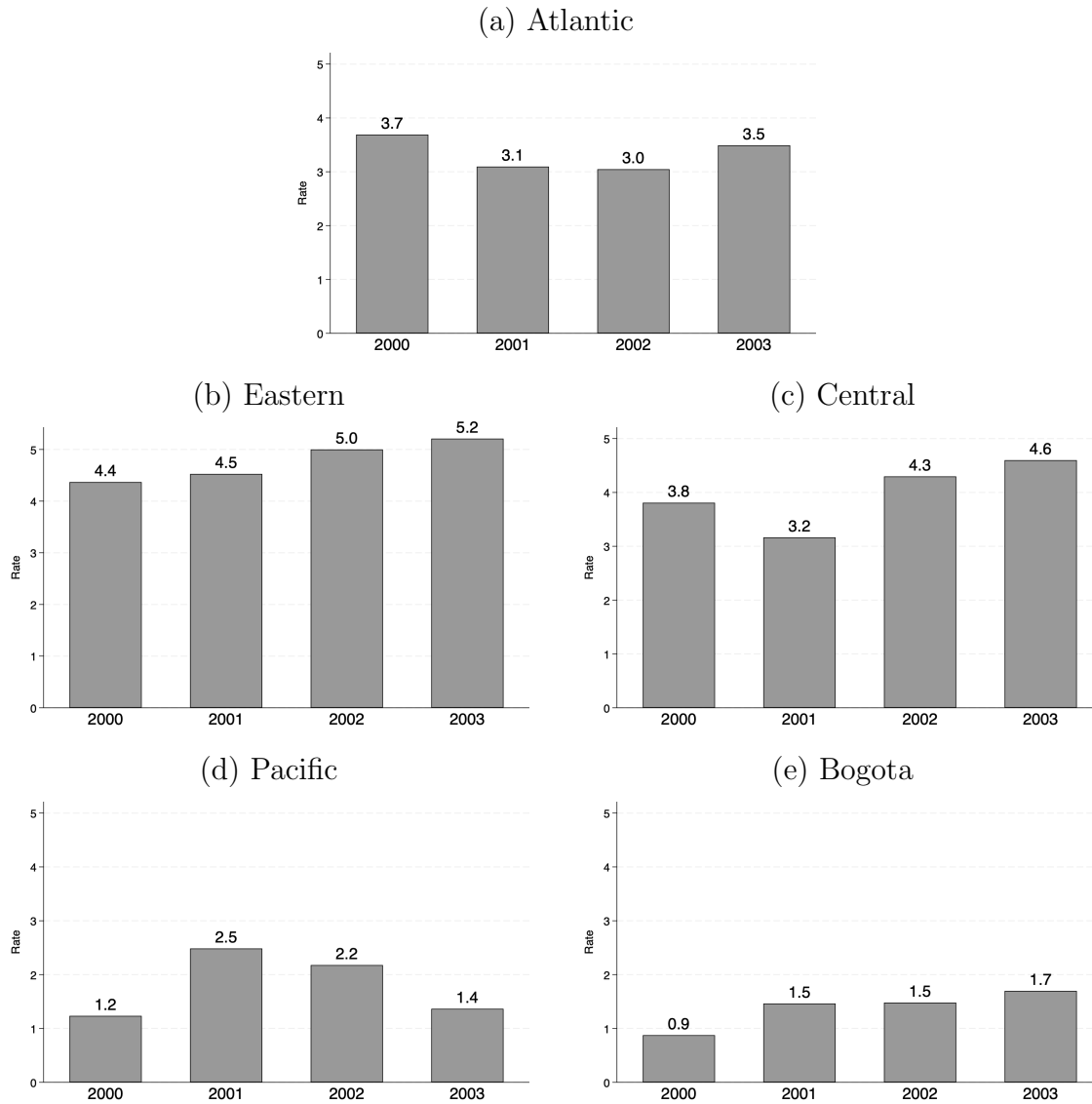
Notes: The table reports estimates from regressions of the leave-one-out proportion of high-achiever peers on cognitive skills. All regressions are limited to schools with at least two classrooms per grade. All models include controls for children's biological sex, age, age squared, and the lagged test 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.

Appendix C

Appendix for Chapter 3

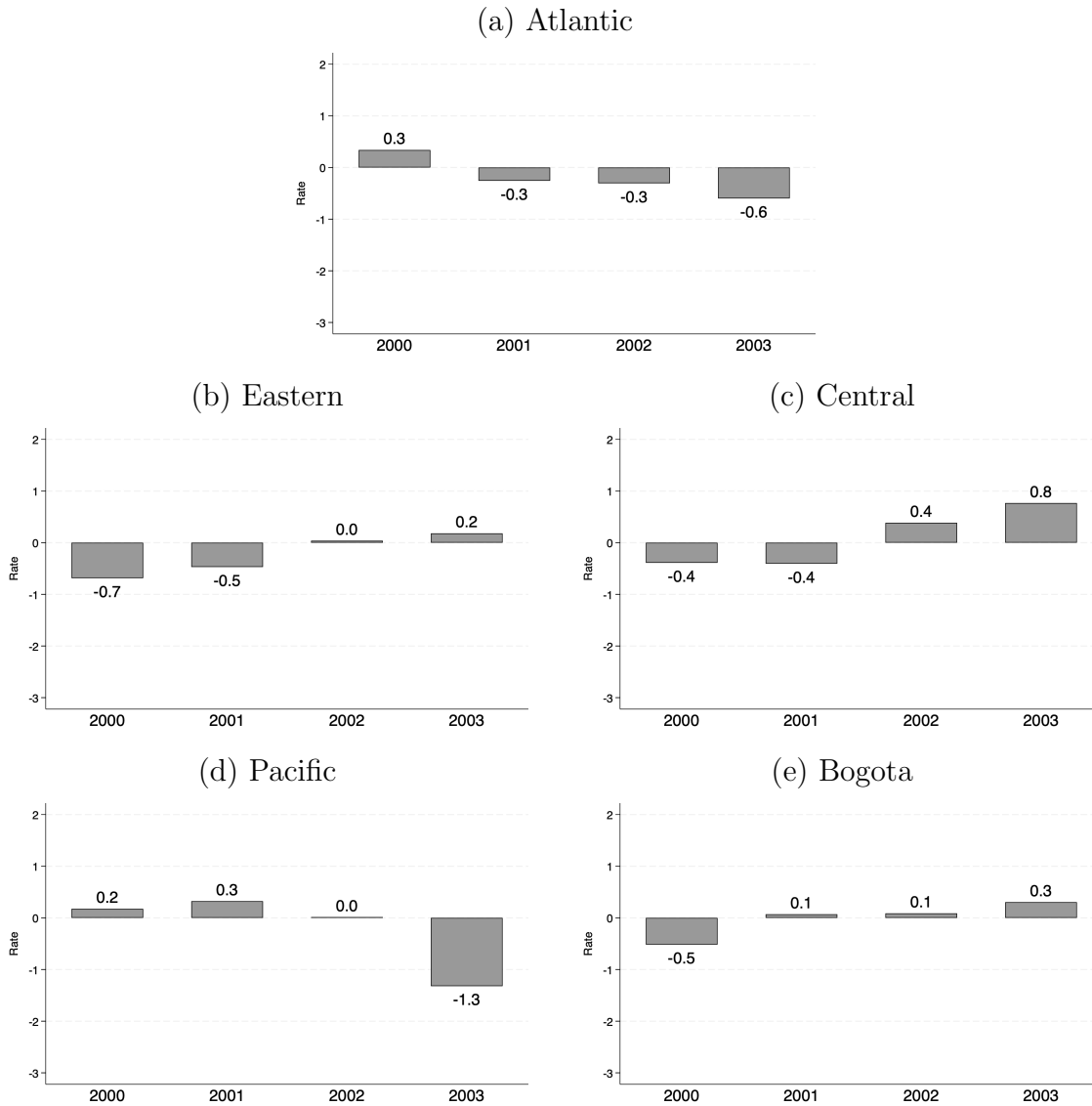
C.1 Supplemental Figures and Tables

Figure C.1: Average Rate of terrorist attacks between 0-9 years old, split by region



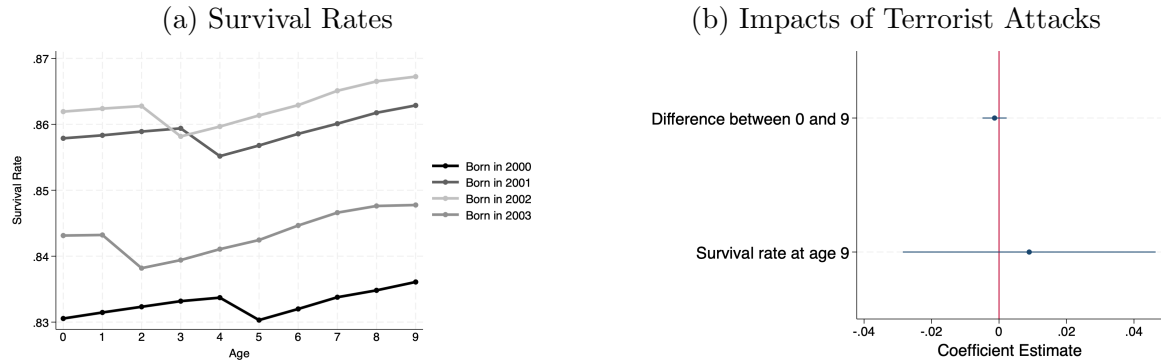
Notes: The figure above shows the average rate of terrorist attacks that a child was exposed to between 0 and 9 years per cohort without accounting for the differences in levels across municipalities. The regions follow the same division as the Colombian Longitudinal Survey.

Figure C.2: Average Rate of terrorist attacks between 0-9 years old, split by region



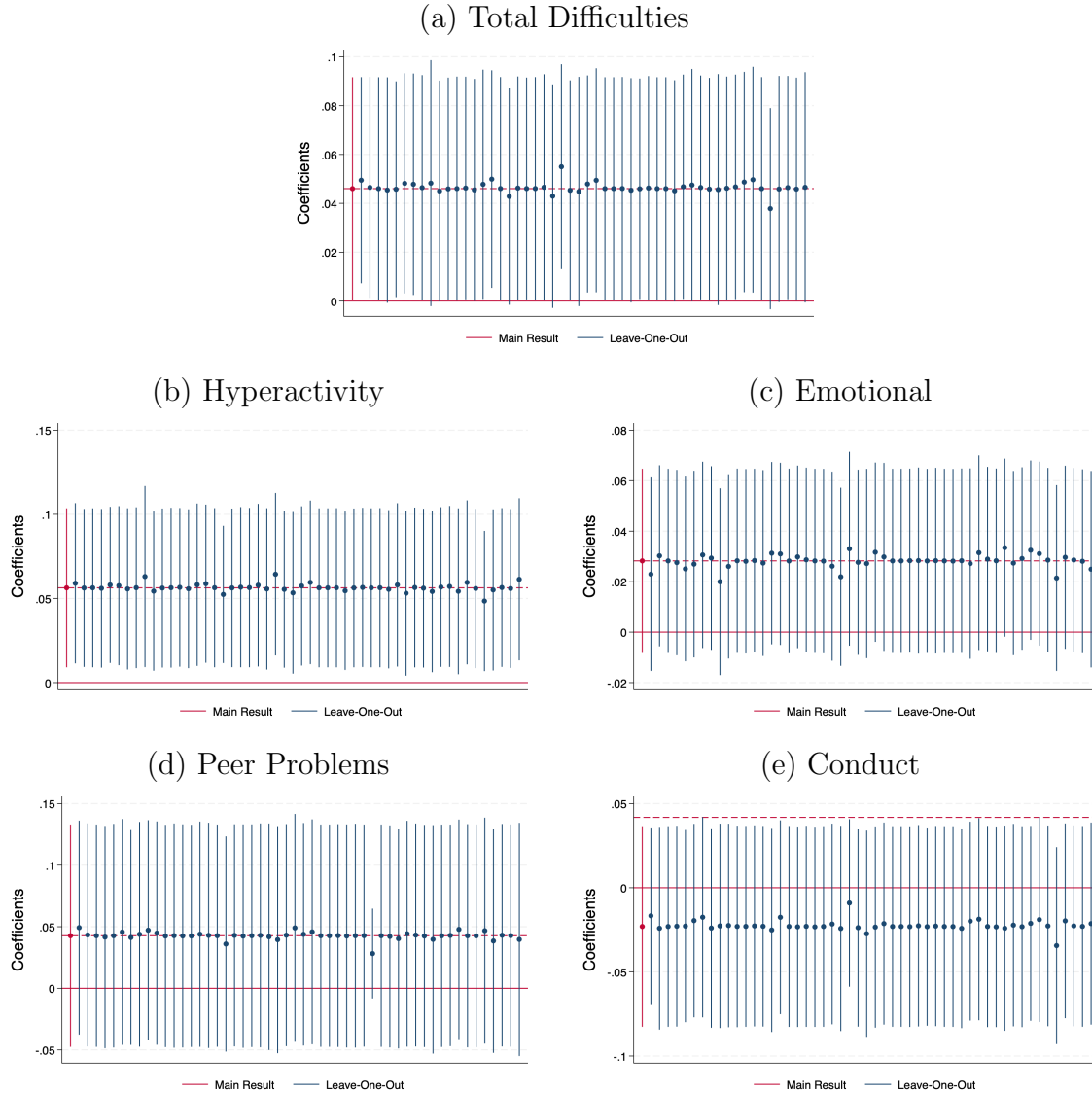
Notes: The figure above shows the demeaned average rate of terrorist attacks that a child was exposed to between 0 and 9 years per cohort accounting for the differences in levels across municipalities. The regions follow the same division as the Colombian Longitudinal Survey.

Figure C.3: Effect of Terrorist Attacks on Survival Rate



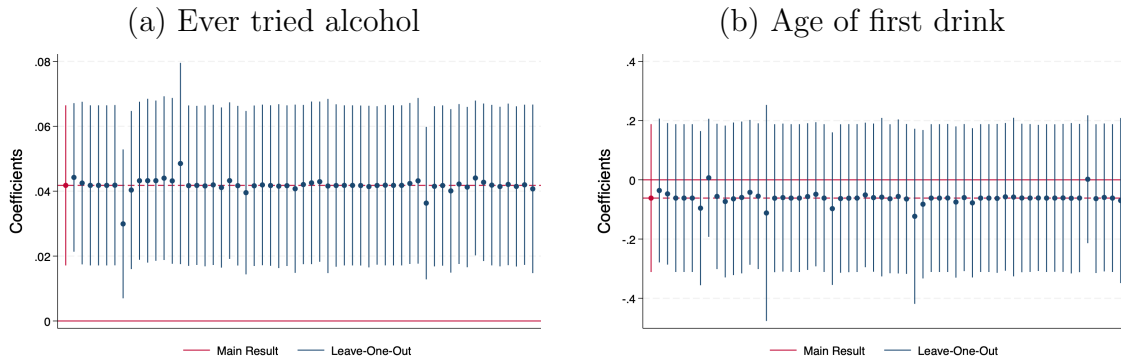
Notes: The figure above shows the impact of the average rate of terrorist attacks per 100,000 inhabitants between 0 and 9 years old on the survival rate of the kids born at the municipality in each cohort. Panel (a) shows the average survival rates for each cohort in the municipalities in the sample. Panel (b) reports estimates from regressions of the rate of terrorist attacks the children in the municipality from a given cohort were exposed to between 0 and 9 years old on the difference between the survival rates at age 9 and age 0, and the survival rate at age 9. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Figure C.4: Leave-one-out Regressions of Socioemotional Skills



Notes: The "Main Result" estimate (on the left side) uses the specification of the even columns of Table 3.4. The other lines use the same specification with one municipality omitted from the sample. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. Results are weighted by children's ELCA weights. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

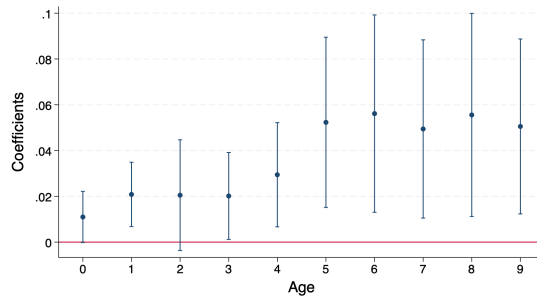
Figure C.5: Leave-one-out Regressions of Alcohol Consumption



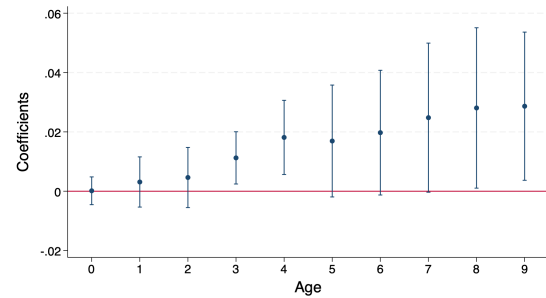
Notes: The "Main Result" estimate (on the left side) uses the specification of the even columns of Table 3.5. The other lines use the same specification with one municipality omitted from the sample. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. Results are weighted by children's ELCA weights. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Figure C.6: Effects of exposure up to certain age

(a) Socioemotional Skills (Total Score)



(b) Ever Tried Alcohol



Notes: The table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to up to a certain age. For example, the estimate for 1 uses the rate between 0 and 1 years old, while the estimate for 7 uses the rate between 0 and 7 years. This figure is a more general version of Tables 3.6 and 3.7. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table C.1: Summary Statistics by cohort

	2000	2001	2002	2003
A. Adolescents characteristics				
Sex (1 = Male)	0.550 (0.499)	0.470 (0.500)	0.516 (0.501)	0.446 (0.499)
Age in years	15.18 (0.388)	14.40 (0.490)	13.38 (0.487)	13 (0)
Years of education	8.061 (1.536)	7.382 (1.461)	6.393 (1.475)	6.069 (0.958)
B. Household characteristics				
Two parent household	0.558 (0.498)	0.550 (0.498)	0.604 (0.490)	0.585 (0.495)
Household size	5.503 (2.396)	5.407 (2.296)	5.495 (2.974)	5.246 (1.817)
Total anual expenditure (in millions)	17.15 (24.29)	16.35 (32.27)	14.78 (13.96)	13.44 (15.00)
C. Caretaker characteristics				
Age	42.52 (8.364)	42.31 (8.445)	41.34 (8.193)	39.46 (7.754)
Years of education	8.913 (4.346)	9.315 (4.336)	8.688 (4.500)	9.433 (4.355)
Participation in social organizations (Number)	0.273 (0.563)	0.218 (0.531)	0.175 (0.507)	0.205 (0.493)
Leader in a social organizations (Number)	0.105 (0.405)	0.0877 (0.326)	0.0706 (0.344)	0.0787 (0.298)

Notes: The table above reports summary statistics of the municipalities and adolescents in the sample by cohort. It includes adolescents' characteristics and those of their households and caretakers. The table only includes the municipalities in the sample of analysis, not all of the country. Adolescent variables are from the 2016 survey, while household and caretaker characteristics are from the baseline (2010).

Table C.2: Summary Statistics by region

	Atlantic	Eastern	Central	Pacific	Bogota
A. Adolescents characteristics					
Sex (1 = Male)	0.530 (0.500)	0.488 (0.501)	0.507 (0.502)	0.492 (0.502)	0.422 (0.496)
Age in years	14.11 (0.928)	14.06 (0.890)	14.09 (0.847)	14 (0.843)	13.79 (0.790)
Years of education	6.837 (1.880)	7.227 (1.465)	6.882 (1.471)	6.992 (1.351)	7.312 (1.215)
B. Household characteristics					
Two parent household	0.570 (0.496)	0.643 (0.480)	0.380 (0.487)	0.574 (0.497)	0.672 (0.471)
Household size	6.134 (2.815)	5.086 (2.699)	5.168 (2.175)	5.082 (1.909)	4.844 (1.471)
Total anual expenditure (in millions)	12.56 (13.08)	19.32 (38.37)	13.06 (13.55)	12.69 (9.566)	22.70 (29.91)
C. Caretaker characteristics					
Age	42.35 (8.787)	40.84 (6.977)	42.35 (8.541)	40.24 (8.465)	41.76 (8.483)
Years of education	8.425 (4.839)	9.245 (4.144)	9.550 (4.246)	8.864 (3.907)	10.02 (4.014)
Participation in social organizations (Number)	0.265 (0.542)	0.205 (0.514)	0.163 (0.464)	0.136 (0.412)	0.228 (0.638)
Leader in a social organizations (Number)	0.0948 (0.345)	0.0800 (0.338)	0.0853 (0.354)	0.0254 (0.158)	0.122 (0.454)

Notes: The table above reports summary statistics of the municipalities and adolescents in the sample by region. It includes adolescents' characteristics and those of their households and caretakers. The table only includes the municipalities in the sample of analysis, not all of the country. Adolescent variables are from the 2016 survey, while household and caretaker characteristics are from the baseline (2010).

Table C.3: Effects of Terrorist Attacks on Smoking

	Own Smoking		Friend's smoking	
	(1)	(2)	(3)	(4)
A. Within-Municipality Cross-Cohort Variation				
Terrorist Attacks 0-9 years old	-0.007 (0.007)	-0.007 (0.006)	0.003 (0.011)	0.010 (0.012)
Terrorist Attacks after 10 years old	-0.004 (0.023)	-0.003 (0.021)	-0.020 (0.016)	-0.006 (0.016)
Municipality FE	Yes	Yes	Yes	Yes
B. Nearest Neighbor Match				
Terrorist Attacks 2000-2016 > Median		0.070*** (0.000)		-0.134*** (0.001)
Municipality FE	Yes	No	Yes	No
Mean of Dependent Variable	0.0926	0.0926	0.376	0.376
Observations	870	870	702	702
Controls	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes

Notes: Panel A in the table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on Smoking Outcomes. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table C.4: Effects of Homicides on Socioemotional Skills

	Total Score	Hyperactivity	Emotional	Peers problems	Conduct
	(1)	(2)	(3)	(4)	(5)
Homicide rate 0-9 years old	0.000 (0.002)	-0.001 (0.002)	0.003* (0.001)	0.001 (0.002)	-0.002 (0.002)
Mean of Dependent Variable	8.811	3.355	1.939	1.300	2.227
Observations	856	856	856	856	856
Controls	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of homicides a child was exposed to between 0 and 9 years old on Socioemotional Skills Scores. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table C.5: Effects of Homicides on Alcohol Consumption

	Ever tried alcohol	Age of first drink
	(1)	(2)
Homicides 0-9 years old	-0.0000 (0.0012)	-0.0028 (0.0045)
Mean of Dependent Variable	0.643	12.84
Observations	870	495
Controls	Yes	Yes
Cohort FE	Yes	Yes
Municipality FE	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of homicides a child was exposed to between 0 and 9 years old on Alcohol Consumption Outcomes. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

C.2 Details on Sample Selection

In this appendix, I cover details of the sample selection process. Recall from section 3.3.1 that the survey started in 2010. The first wave included over 5,000 households (strata 1 to 4). At baseline, the target population consisted of the head of the household, their partner, and any children under nine years old living in the household. Regardless of whether they subsequently left the household, these individuals were followed across the three waves in which the survey was conducted until 2016. Indeed, in 2016, 89% of the original households were interviewed.

The baseline survey also collected basic information from an individual not selected to be in the target population. For subsequent waves, any new household member would also not be considered in the target population and would only answer a small questionnaire. For example, if, in 2013, a kid and the head of the household moved with her grandmother, then the kid and the head of the household would answer the complete questionnaire while the grandmother would only answer the basic one. However, if the kid moved with her grandmother, both would answer the complete questionnaire.

In 2016, the survey included risky health behaviors questions and the Strengths and Difficulties Questionnaire (SDQ), which was only for the children selected as the target population in 2010 and were older than ten at the moment of the survey in 2016. Therefore, the information is only available for adolescents aged 10 to 16.

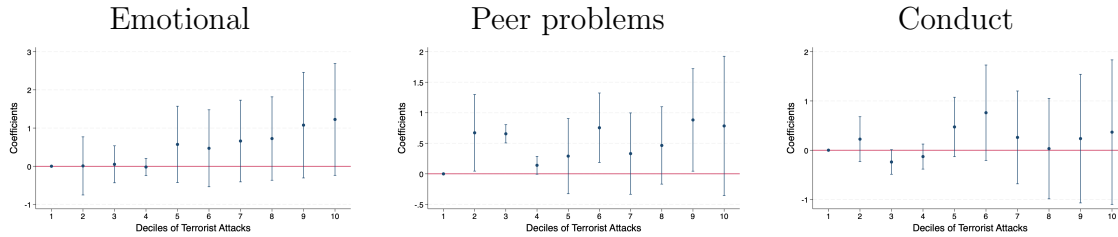
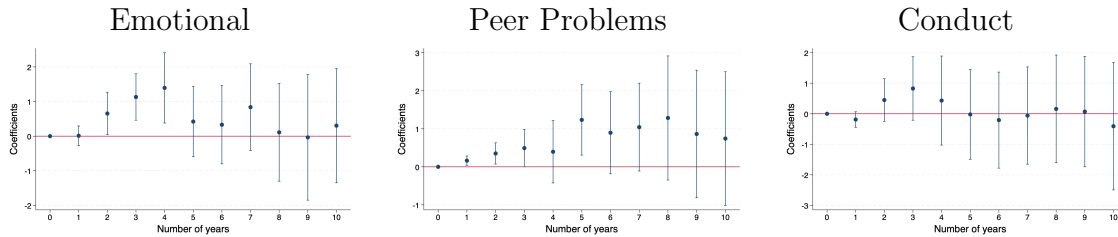
Based on the evidence for Colombia, children typically start drinking after 12 years old (de la Espriella Guerrero et al., 2016). Therefore, I restrict the sample to children aged 13 to 16. The final sample consists of 870 children born between 2000 and 2003. The sample includes data from 832 households out of the more than 5,000 included in the survey, as only they have children in the specified age range.

C.3 Extra estimates on Socioemotional Skills Components

As explained in Section 3.5, there are no significant effects on three of the components of the socioemotional skills score: emotional, peer problems, and conduct. However, this section presents all analyses conducted for the total difficulties and hyperactivity scores. This section includes the following figures and tables:

- Non-linear effects of terrorist attacks as presented in section 3.5.2
- Effects of terrorist attacks by childhood periods as presented in section 3.5.3
- Robustness and alternative specification checks as presented in section 3.5.4
- Heterogeneous effects as presented in section 3.5.5
- Placebo effects as presented in section 3.5.6

Figure C.7: Non-Linear Impacts of Terrorist Attacks on Socioemotional Skills Components

Panel A. Deciles of Rate of Terrorist Attacks**Panel B. Number of years exposed to terrorist attacks**

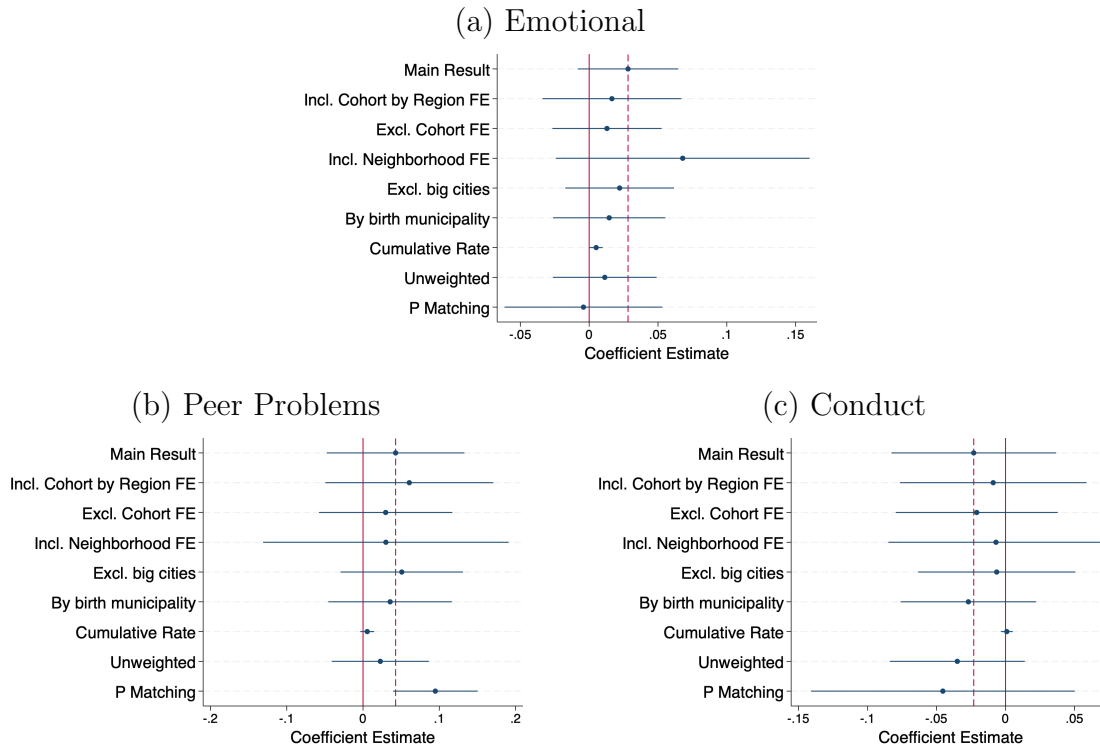
Notes: Panel A in the figure above reports estimates from regressions of the deciles of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on the remaining Socioemotional Skills Components Scores. The first decile is the omitted category. Panel B in the figure above reports estimates from regressions of the number of years with strictly greater than 0 terrorist attacks exposure between 0 and 9 years old on the remaining Socioemotional Skills Components Scores. 0 years of exposure is the omitted category. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. Results are weighted by children's ELCA weights. Vertical bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table C.6: Effects of Terrorist Attacks on Socioemotional Skills Components by Childhood periods

	Emotional		Peer problems		Conduct	
	(1)	(2)	(3)	(4)	(5)	(6)
Terrorist Attacks 0-5 years old	0.066*** (0.017)	0.051** (0.022)	0.039 (0.029)	0.037 (0.030)	0.032 (0.023)	0.021 (0.021)
Terrorist Attacks 6-9 years old	-0.063*** (0.021)	-0.055** (0.025)	-0.013 (0.028)	-0.005 (0.032)	-0.099*** (0.025)	-0.078*** (0.022)
Terrorist Attacks after 10 years old	-0.101* (0.053)	-0.105* (0.052)	-0.028 (0.058)	-0.013 (0.069)	-0.206*** (0.055)	-0.174*** (0.049)
Mean of Dependent Variable	1.939	1.939	1.300	1.300	2.227	2.227
Observations	856	856	856	856	856	856
Controls	No	Yes	No	Yes	No	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes

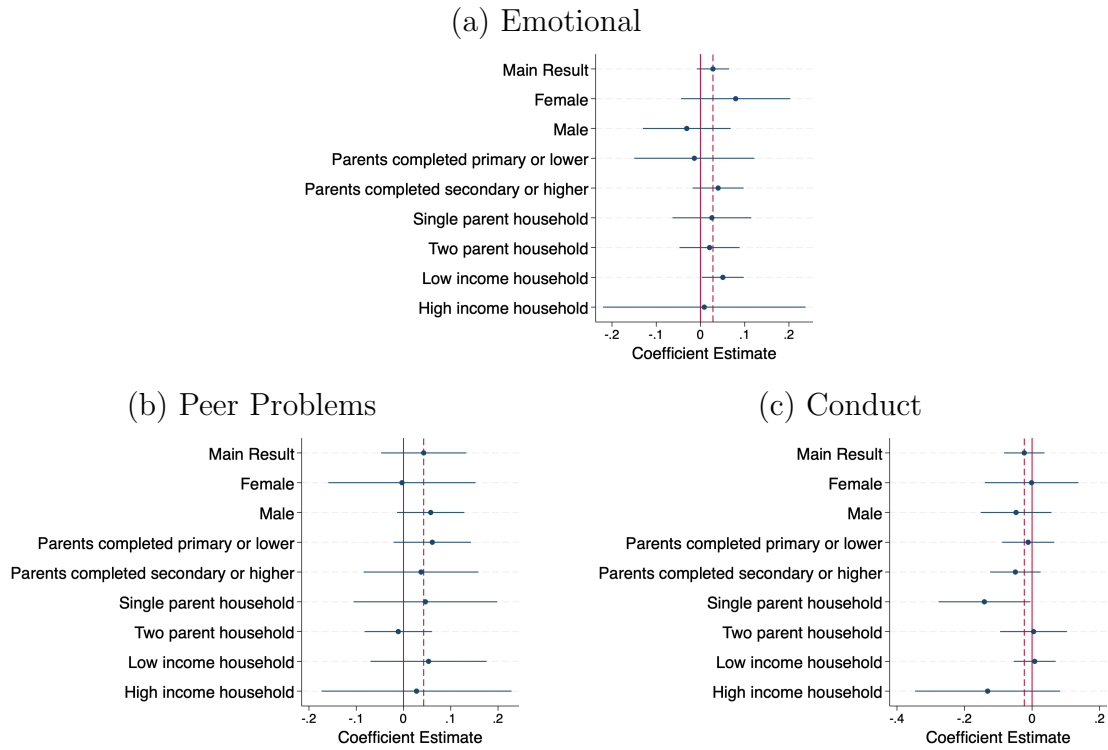
Notes: The table above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 5 years old, and between 6 and 9 years old, separately, on the remaining Socioemotional Skills Components Scores. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Figure C.8: Robustness checks on Socioemotional Skills Components



Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on remaining Socioemotional Skills Components Scores for different specifications. Results are weighted by children's ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level. The "Main Result" estimate (at the top) uses the specification from the even columns of Table 3.4. All other estimates are variations on the baseline model. Estimates 2-4 vary the set of fixed effects. Estimate 5 restricts the sample to small municipalities only, and estimate 6 assigns the rate of terrorist attacks to the birth municipality instead of the residency municipality. Estimate 7 uses the cumulative rate of terrorist attacks between 0 and 9, instead of the average rate, while estimate 8 indicates no weights are used in the regression. Finally, estimate 9 uses a probability matching estimation.

Figure C.9: Heterogeneous Impacts on Socioemotional Skills Components



Notes: The figure above reports estimates from regressions of the rate of terrorist attacks a child was exposed to between 0 and 9 years old on the remaining Socioemotional Skills Components Scores separated based on characteristics of the children, their parents, and their household. Results are weighted by children's ELCA weights. Horizontal bars represent 95% confidence intervals. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Table C.7: Placebo Effects of Terrorist Attacks on Socioemotional Skills

	Emotional	Peers Problems	Conduct
	(1)	(2)	(3)
Terrorist Attacks 2-7 years before	-0.050* (0.025)	-0.005 (0.025)	0.010 (0.014)
Mean of Dependent Variable	1.939	1.300	2.227
Observations	856	856	856
Controls	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes

Notes: The table above reports estimates from regressions of the rate of terrorist attacks between 2 and 7 years the child was born on Socioemotional Skills Scores. All models include controls for children's biological sex, education, order within the children of the household, age of caretaker, education of the caretaker, whether the household is a two-parent in 2010, annual logarithm of consumption in 2010, and household size in 2010. *** indicates significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level. Results are weighted by children's ELCA weights. Standard errors are corrected for heteroskedasticity and are clustered at the municipality level.

Bibliography

- Abadie, A. and Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132.
- Akbulut-Yuksel, M. (2014). Children of War: The Long-run Effects of Large-scale Physical Destruction and Warfare on Children. *Journal of Human Resources*, 49(3):634–662.
- Akresh, R., Lucchetti, L., and Thirumurthy, H. (2012). Wars and Child Health: Evidence from the Eritrean-Ethiopian Conflict. *Journal of Development Economics*, 99(2):330–340.
- Allais, O., Fagherazzi, G., and Mink, J. (2021). The Long-run Effects of War on Health: Evidence from World War II in France. *Social Science and Medicine*, 276.
- Almond, D. and Currie, J. (2011). Human Capital Development before Age Five. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier.
- Almond, D., Currie, J., and Duque, V. (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature*, 56:1360–1446.
- Angrist, J. D. (2014). The Perils of Peer Effects. *Labour Economics*, 30:98.
- Arain, M., Haque, M., Johal, L., Mathur, P., Nel, W., Rais, A., Sandhu, R., and Sharma, S. (2013). Maturation of the Adolescent Brain. *Neuropsychiatric Disease and Treatment*, 9:449–461.
- 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.
- Araujo, M. D., Cruz-Aguayo, Y., and Heineck, G. (2025). Using Screening Assessments to Recruit Effective Teachers: Experimental Evidence from Ecuador. *IDB Working Paper 1677*.

BIBLIOGRAPHY

- Aronow, P. M., Samii, C., and Assenova, V. A. (2017). Cluster-Robust Variance Estimation for Dyadic Data. *Political Analysis*, 23(4):564–577.
- Asselmann, E., Wittchen, H. U., Lieb, R., Höfler, M., and Beesdo-Baum, K. (2016). Does Low Coping Efficacy Mediate the Association between Negative Life Events and Incident Psychopathology? A Prospective-Longitudinal Community Study among Adolescents and Young Adults. *Epidemiology and Psychiatric Sciences*, 25(2):171–180.
- Babcock, P., Bedard, K., Charness, G., Hartman, J., and Royer, H. (2015). Letting Down the Team? Social Effects of Team Incentives. *Journal of the European Economic Association*, 13(5):841–870.
- Bailey, M. J., Hoynes, H., Rossin-Slater, M., and Walker, R. (2024). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence From the Food Stamps Program. *Review of Economic Studies*, 91.
- Bailey, M. J., Hoynes, H. W., Rossin-Slater, M., and Walker, R. (2020). Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program. *NBER Working Paper Series 26942*.
- Balestra, S., Sallin, A., and Wolter, S. C. (2023). High-Ability Influencers? The Heterogeneous Effects of Gifted Classmates. *Journal of Human Resources*, 58(2).
- Bandiera, O., Barankay, I., and Rasul, I. (2010). Social Incentives in the Workplace. *Review of Economic Studies*, 77(2).
- Barrios-Fernández, A. (2022). Neighbors’ Effects on University Enrollment. *American Economic Journal: Applied Economics*, 14(3):30.
- Bassi, M., Meghir, C., and Reynoso, A. (2020). Education Quality and Teaching Practices. *Economic Journal*, 130(631):1937–1965.
- Bastian, J. and Micheltore, K. (2018). The Long-term Impact of the Earned Income Tax Credit on Children’s Education and Employment Outcomes. *Journal of Labor Economics*, 36(4):1127–1163.
- Bayer, P., Ross, S. L., and Topa, G. (2008). Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, 116(6):1150–1196.
- 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.

BIBLIOGRAPHY

- Bharadwaj, P., Eberhard, J. P., and Neilson, C. A. (2018). Health at Birth, Parental Investments, and Academic Outcomes. *Journal of Labor Economics*, 36(2):349–394.
- 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.
- Billings, S. B., Deming, D. J., and Rockoff, J. (2014). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg. *Quarterly Journal of Economics*, 129(1):435–476.
- 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.
- Blattman, C. and Annan, J. (2010). The Consequences of Child Soldiering. *Review of Economics and Statistics*, 92(4):882–898.
- Borghans, L., ter Weel, B., and Weinberg, B. A. (2008). Interpersonal Styles and Labor Market Outcomes. *Journal of Human Resources*, 43(4):815–858.
- Brade, R. (2025). Short-Term Events, Long-Term Friends? Freshman Orientation Peers and Academic Performance. *Journal of Human Resources*, 60(4):1405–1434.
- Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of Peer Effects Through Social Networks. *Journal of Econometrics*, 150(1):415–55.
- Brodeur, A. and Yousaf, H. (2022). On the Economic Consequences of Mass Shootings. *The Review of Economics and Statistics*, pages 1–43.
- Burke, N. J., Hellman, J. L., Scott, B. G., Weems, C. F., and Carrion, V. G. (2011). The Impact of Adverse Childhood Experiences on an Urban Pediatric Population. *Child Abuse and Neglect*, 35(6):408–413.
- Busso, M. and Frisanchi, V. (2021). Good Peers have Asymmetric Gendered Effects on Female Educational Outcomes: Experimental Evidence from Mexico. *Journal of Economic Behavior and Organization*, 189.
- Camacho, A. (2008). Stress and Birth Weight: Evidence from Terrorist Attacks. *American Economic Review*, 98(2):511–515.
- Cameron, A. C. and Miller, D. (2014). Robust Inference for Dyadic Data.
- 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.

BIBLIOGRAPHY

- Carneiro, P., Cruz-Aguayo, Y., Salvati, F., and Schady, N. (2025a). Disruption in the Classroom: Experimental Evidence from Ecuador. *IDB Working Paper 1688*.
- Carneiro, P., Cruz-Aguayo, Y., Salvati, F., and Schady, N. (2025b). The Effect of Classroom Rank on Learning throughout Elementary School: Experimental Evidence from Ecuador. *Journal of Labor Economics*.
- Carneiro, P., Cruz-Aguayo, Y., and Schady, N. (2020). Mothers, Teachers, Peers, and the Gender Gap in Early Math Achievement. *IDB Working Paper 00939*.
- Carrell, 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.
- 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.
- Cawley, J. and Ruhm, C. J. (2011). The Economics of Risky Health Behaviors. In Pauly, M. V., McGuire, T. G., and Barros, P. P., editors, *Handbook of Health Economics*, volume 2, pages 95–199. Elsevier.
- Centro Nacional de Memoria Histórica (2013). *¡Basta ya! Colombia: Memorias de Guerra y Dignidad*. Bogotá:, Imprenta Nacional.
- Chassin, L., Presson, C. C., Rose, J. S., and Sherman, S. J. (1996). The Natural History of Cigarette Smoking from Adolescence to Adulthood: Demographic Predictors of Continuity and Change. *Health Psychology*, 15(6):478–484.
- Chen, S. and Hu, Z. (2024). How Competition Shapes Peer Effects: Evidence from a University in China. *The Review of Economics and Statistics*, pages 1–27.
- Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added 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.
- Chetty, R., Hendren, N., and Katz, L. F. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902.

BIBLIOGRAPHY

- Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022). Social Capital I: Measurement and Associations with Economic Mobility. *Nature*, 608(7921):108.
- Coleman, J. S. (1958). Relational Analysis: The Study of Social Organizations with Survey Methods. *Human Organization*, 17(4):28–36.
- Conti, G. and Heckman, J. (2012). The Economics of Child Well-Being. *NBER Working Paper No. 18466*.
- Cook, P. J. and Moore, M. J. (2001). Environment and Persistence in Youthful Drinking Patterns. In Gruber, J., editor, *Risky Behavior among Youths: An Economic Analysis*, chapter 8, pages 375–438. University of Chicago Press, Chicago, IL.
- Cools, A., Fernández, R., and Patacchini, E. (2022). The Asymmetric Gender Effects of High Flyers. *Labour Economics*, 79.
- Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer Effects in the Workplace. *American Economic Review*, 107(2):425–56.
- Cunha, F. and Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, 97(2):31–47.
- Currarini, S., Jackson, M. O., and Pin, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4):1003–1045.
- Currie, J. (2001). Early Childhood Education Programs. *Journal of Economic Perspectives*, 15(2):213–238.
- de Gendre, A. and Salamanca, N. (2020). On the Mechanisms of Ability Peer Effects. *IZA DP No. 13938*.
- de la Espriella Guerrero, R. A., Rodriguez, V., Rincón, C. J., Morales, D. C., Rodríguez, S. J. P., and Gómez-Restrepo, C. (2016). Consumo de Alcohol en la Población Colombiana. Encuesta Nacional de Salud Mental 2015. *Revista Colombiana de Psiquiatria*, 45(S1):76–88.
- de Roux, N. and Riehl, E. (2022). Do College Students Benefit from Placement into Higher-achieving Classes? *Journal of Public Economics*, 210.
- Deming, D. (2009). Early Childhood Intervention and Life-cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, 1(3):111–134.

BIBLIOGRAPHY

- Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, 132(4):1593–1640.
- Deza, M. (2015). Is There a Stepping Stone Effect in Drug Use? Separating State Dependence from Unobserved Heterogeneity within and between Illicit Drugs. *Journal of Econometrics*, 184(1):193–207.
- Di Maio, M. and Sciabolazza, V. L. (2021). Conflict Exposure and Health: Evidence from the Gaza Strip. *Health Economics*, 30(9).
- Dube, S. R., Miller, J. W., Brown, D. W., Giles, W. H., Felitti, V. J., Dong, M., and Anda, R. F. (2006). Adverse Childhood Experiences and the Association with ever using Alcohol and Initiating Alcohol Use During Adolescence. *Journal of Adolescent Health*, 38(4):1–10.
- 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.
- Ellsberg, M., Heise, L., Peña, R., Agurto, S., and Winkvist, A. (2001). Researching Domestic Violence against Women: Methodological and Ethical Considerations. *Studies in Family Planning*, 32(1):1–16.
- 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.
- Fernández, M., Ibáñez, A. M., and Peña, X. (2014). Adjusting the Labour Supply to Mitigate Violent Shocks: Evidence from Rural Colombia. *Journal of Development Studies*, 50(8):1135–1155.
- Figlio, D., Guryan, J., Karbownik, K., and Roth, J. (2014). The Effects of Poor Neonatal Health on Children’s Cognitive Development. *American Economic Review*, 104(12):3921–3955.
- Fletcher, J. M. and Ross, S. L. (2018). Estimating the Effects of Friends on Health Behaviors of Adolescents. *Health Economics*, 27(10):1450–1483.
- Fletcher, J. M., Ross, S. L., and Zhang, Y. (2020). The Consequences of Friendships: Evidence on the Effect of Social Relationships in School on Academic Achievement. *Journal of Urban Economics*, 116.

BIBLIOGRAPHY

- Friedman, A. S. (2020). Smoking to Cope: Addictive Behavior as a Response to Mental Distress. *Journal of Health Economics*, 72.
- Fuertes-Segura, N., Cruz-Aguayo, Y., and Echeverri, C. (2025). The Asymmetric Effects of High Achiever Peers: Experimental Evidence from Ecuador. *SSRN Electronic Journal*.
- Garrett, A., Cohen, J. A., Zack, S., Carrion, V., Jo, B., Blader, J., Rodriguez, A., Vanasse, T. J., Reiss, A. L., and Agras, W. S. (2019). Longitudinal Changes in Brain Function Associated with Symptom Improvement in Youth with PTSD. *Journal of Psychiatric Research*, 114:161–169.
- 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*.
- Giedd, J. N. (2004). Structural Magnetic Resonance Imaging of the Adolescent Brain. *Annals of the New York Academy of Sciences*, 1021(1):77–85.
- Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social Networks and the Identification of Peer Effects. *Journal of Business and Economic Statistics*, 31(3):253–264.
- Golsteyn, B. H., Non, A., and Zölitz, U. (2021). The Impact of Peer Personality on Academic Achievement. *Journal of Political Economy*, 129(4):1052.
- Goodreau, S. M., Kitts, J. A., and Morris, M. (2009). Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks. *Demography*, 46(1):103–125.
- Gould, E. D., Lavy, V., and Paserman, M. D. (2011). Sixty Years after the Magic Carpet Ride: The Long-run Effect of the Early Childhood Environment on Social and Economic Outcomes. *Review of Economic Studies*, 78(3):938–973.
- Graham, B. S. (2017). An Econometric Model of Network Formation With Degree Heterogeneity. *Econometrica*, 85(4):1033–1063.
- Graham, B. S. (2020). Network Data. *Handbook of Econometrics*, 7.
- Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380.
- Grossman, D. and Khalil, U. (2022). Neighborhood Crime and Infant Health. *Journal of Urban Economics*, 130.
- Grossman, M. and Markowitz, S. (2005). I Did What Last Night? Adolescent Risky Sexual Behaviors and Substance Abuse. *Eastern Economic Journal*, 31(3):383–405.

BIBLIOGRAPHY

- Gruber, J. and Zinman, J. (2001). Youth Smoking in the United States. In Gruber, J., editor, *Risky Behavior among Youths: An Economic Analysis*, chapter 2, pages 69–120. University of Chicago Press, Chicago, IL.
- Guryan, J., Kroft, K., and Notowidigdo, M. J. (2009). Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68.
- 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. (2000). Policies to Foster Human Capital. *Research in Economics*, 54(1):3–56.
- 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. and Rubinstein, Y. (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *American Economic Review*, 91(2):145–149.
- 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.
- Hellerstein, J. K., Mcinerney, M., and Neumark, D. (2011). Neighbors and Coworkers: The Importance of Residential Labor Market Networks. *Journal of Labor Economics*, 29(4):659–695.
- Hensvik, L. and Skans, O. N. (2016). Social Networks, Employee Selection, and Labor Market Outcomes. *Journal of Labor Economics*, 34(4):825–867.
- 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*.
- Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review*, 106(4):903–934.
- Institute for Economics and Peace (2023). Global Terrorism Index. Retrieved from <https://www.economicsandpeace.org/reports/>.

BIBLIOGRAPHY

- Jackson, M. O., Nei, S., Snowberg, E., and Yariv, L. (2025). The Dynamics of Networks and Homophily. *SSRN Electronic Journal*.
- Jacob, B. A., Lefgren, L., and Sims, D. P. (2010). The Persistence of Teacher-induced Learning. *Journal of Human Resources*, 45(4):915–943.
- Jacobson, M. and Royer, H. (2011). Aftershocks: The Impact of Clinic Violence on Abortion Services. *American Economic Journal: Applied Economics*, 3(1):189–223.
- Jochmans, K. (2023). Testing Random Assignment to Peer Groups. *Journal of Applied Econometrics*, 38(3):321–333.
- Johnsson, I. and Moon, H. R. (2021). Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach. *Review of Economics and Statistics*, 103(2):328–345.
- Jürges, H., Stella, L., Hallaq, S., and Schwarz, A. (2022). Cohort at Risk: Long-term Consequences of Conflict for Child School Achievement. *Journal of Population Economics*, 35(1):1–43.
- Koppensteiner, M. F. and Menezes, L. (2021). Violence and Human Capital Investments. *Journal of Labor Economics*, 39(3):787–823.
- Kreif, N., Mirelman, A., Suhrcke, M., Buitrago, G., and Moreno-Serra, R. (2022). The Impact of Civil Conflict on Child Health: Evidence from Colombia. *Economics and Human Biology*, 44.
- Lavy, V. and Sand, E. (2019). The Effect of Social Networks on Students’ Academic and Non-Cognitive Behavioural Outcomes: Evidence from Conditional Random Assignment of Friends in School. *Economic Journal*, 129(617):439–480.
- 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.
- Levine, P. B. (2001). The Sexual Activity and Birth-control use of American Teenagers. In Gruber, J., editor, *Risky Behavior among Youths: An Economic Analysis*, chapter 4, pages 167–218. University of Chicago Press, Chicago, IL.
- León, G. (2012). Civil Conflict and Human Capital Accumulation: The Long-term Effects of Political Violence in Perú. *Journal of Human Resources*, 47(4):991–1022.
- Lindqvist, E. and Vestman, R. (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics*, 3(1):101–128.

BIBLIOGRAPHY

- Ludwig, J. and Miller, D. L. (2007). Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *Quarterly Journal of Economics*, 122(1):159–208.
- Manian, S. (2021). Conflict and Risky Health Behavior: Evidence from Mexico's Drug War. *Journal of Development Economics*, 148.
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60(3):531–542.
- Marshall, M. G. (2020). Major Episodes of Political Violence 1946-2019. Retrieved from <https://www.systemicpeace.org/warlist/warlist.htm>.
- 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.
- Mezza, A. and Buchinsky, M. (2021). Illegal Drugs, Education, and Labor Market Outcomes. *Journal of Econometrics*, 223(2):454–484.
- Minoiu, C. and Shemyakina, O. N. (2014). Armed Conflict, Household Victimization, and Child Health in Côte d'Ivoire. *Journal of Development Economics*, 108:237–255.
- 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.
- Moody, J. (2001). Race, School Integration, and Friendship Segregation in America. *American Journal of Sociology*, 107(3):679–716.
- 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.
- National Scientific Council on the Developing Child (2010). Persistent Fear and Anxiety Can Affect Young Children's Learning and Development: Working Paper No. 9. Retrieved from www.developingchild.harvard.edu.
- National Scientific Council on the Developing Child (2011). Building the Brain's Air Traffic Control System: How Early Experiences Shape the Development of Executive Function: Working Paper No. 11. Retrieved from www.developingchild.harvard.edu.

BIBLIOGRAPHY

- National Scientific Council on the Developing Child (2015). Excessive Stress Disrupts the Architecture of the Developing Brain: Working Paper 3. Retrieved from www.developingchild.harvard.edu.
- 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.
- Park, S. (2019). Socializing at Work: Evidence from a Field Experiment with Manufacturing Workers. *American Economic Journal: Applied Economics*, 11(3):424–455.
- Pearson, M., Steglich, C. E., and Snijders, T. (2006). Homophily and Assimilation Among Sport-active Adolescent Substance Users. *Connections*, 27(1):47–63.
- Pianta, R., Paro, K. L., and Hamre, B. (2015). *Classroom Assessment Scoring System—CLASS*. Brookes, Baltimore, MD.
- Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47(5):651–660.
- Ramirez Muñoz, J. (2015). Exposición Prenatal a la Violencia: Efectos a Largo Plazo sobre el Desempeño Escolar. *Documento CEDE 11*.
- Rees, D. I., Argys, L. M., and Averett, S. L. (2001). New Evidence on the Relationship between Substance Use and Adolescent Sexual Behavior. *Journal of Health Economics*, 20(5):835–845.
- Renna, F. (2008). Teens’ Alcohol Consumption and Schooling. *Economics of Education Review*, 27(1):69–78.
- Rodriguez, C. and Sanchez, F. (2009). Armed Conflict Exposure, Human Capital Investments and Child Labor: Evidence from Colombia. *Documento CEDE 5*.
- Rodriguez, C. and Sanchez, F. (2010). Books and Guns: The Quality of Schools in Conflict Zones. *Documento CEDE 38*.
- Rothstein, J. (2010). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics*, 125(1):175–214.
- Rury, D. (2025). Putting the K in Rank: How Kindergarten Classrooms Impact Short and Long-Run Outcomes. *Applied Economics*, pages 1–15.
- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116(2):681–704.

BIBLIOGRAPHY

- 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.
- Santavirta, T. and Sarzosa, M. (2025). Effects of Disruptive Peers in Endogenous Social Networks. *Quantitative Economics*.
- Shany, A. (2023). Too Scared for School? Effects of Terrorism on Students’ Achievement. *Journal of Human Resources*.
- Shemyakina, O. N. (2021). Political Violence and Child Health: Results from Zimbabwe. *Economics and Human Biology*, 42.
- Shonkoff, J. P., Garner, A. S., Siegel, B. S., Dobbins, M. I., Earls, M. F., McGuinn, L., Pascoe, J., Wood, D. L., High, P. C., Donoghue, E., Fussell, J. J., Gleason, M. M., Jaudes, P. K., Jones, V. F., Rubin, D. M., Schulte, E. E., Macias, M. M., Bridgemohan, C., Fussell, J., Goldson, E., McGuinn, L. J., Weitzman, C., and Wegner, L. M. (2012). The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics*, 129(1):232–246.
- Shrum, W., Cheek, N. H., and Hunter, S. M. (1988). Friendship in School: Gender and Racial Homophily. *Sociology of Education*, 61(4):227–239.
- 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).
- Sánchez-Ariza, J., Cuartas, J., and Moya, A. (2023). The Mental Health of Caregivers and Young Children in Conflict-Affected Settings. *AEA Papers and Proceedings*, 113.
- Verbrugge, L. M. (1977). The Structure of Adult Friendship Choices. *Social Forces*, 56(2):576–597.
- Villeval, M. C. (2020). Performance Feedback and Peer Effects. In Zimmermann, K. F., editor, *Handbook of Labor, Human Resources and Population Economics*. Springer, Cham.
- Winters, K. C. and Arria, A. (2011). Adolescent Brain Development and Drugs. *The Prevention Researcher*, 18(2):21–24.

BIBLIOGRAPHY

- 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.
- Zhu, M. (2025). Beyond the “Old Boys’ Network”: Social Networks and Job Finding at Community Colleges. *Journal of Human Resources*, 60(4):1405–1434.
- Zárate, R. A. (2023). Uncovering Peer Effects in Social and Academic Skills. *American Economic Journal: Applied Economics*, 15(3).