Effective Families or Effective Schools? Experimental Evidence on Fostering Children's Numeracy*

Samuel Berlinski Michele Giannola Alessandro Toppeta

January 28, 2025

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CITE OR CIRCULATE WITHOUT PERMISSION

Abstract

This paper studies if high-quality primary education can serve as an equalizer for children exposed to diverse rearing practices at home in Colombia. Building on well-established literatures demonstrating that parenting interventions and teacher professional development programs can effectively support children's learning, we randomly assign children to a sequence of treatment combinations: only a parenting intervention, only a teacher intervention, both interventions, or no intervention. The two treatments, both individually and together, have a similar effect on children's standardized math scores, with an impact of 0.20 standard deviations compared to receiving no interventions. We find no complementarities between the two interventions, suggesting that effective schools can compensate for inequalities in family rearing practices.

*Berlinski: Inter-American Development Bank. Email: samuelb@iadb.org. Giannola: University of Naples Federico II, CSEF and the Institute for Fiscal Studies. Email: michele.giannola@unina.it. Toppeta: SOFI, Stockholm University. Email: alessandro.toppeta@sofi.su.se. We would like to thank Fundacion Luker, PREST and STEM Academia for their contribution in implementing the interventions. We thank Susanna Loeb, Greg Veramendi, Erich Battistin, Henning Hermes, Philipp Lergetporer, Miriam Wust and seminar participants at Williams College, IFAU (Uppsala), SOFI (Stockholm University), IDB, CESifo Venice Summer Institute on Funding Education, UCL, LSE, Workshop on Early Investments, Family Well-Being, and Child Development (UCopnaghen/VIVE), Advances in Field Experiments (AFE), TUM Workshop on Field Experiments in Economics and Business for their useful comments. Isabel Avila provided outstanding research assistantship. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

1 Introduction

In the 19th century, American educator Horace Mann championed the concept of formal schooling as the "great equalizer of the conditions" (Mann 1848). He advocated for public education to be universally accessible and equitable, viewing it as a catalyst for social harmony and poverty reduction. However, a recent literature underscores the significance of early childhood in the long-term development of human capital, prompting questions about whether efforts to equalize conditions for children should commence much earlier (Carneiro & Heckman 2002, 2003).

Research indicates, that prior to the beginning of formal schooling, children should be engaged in a playful learning environment, exposed to rich linguistic experiences, and given opportunities to explore their interests and curiosity (Currie & Almond 2011, Almond et al. 2018). Unfortunately, parents from disadvantaged backgrounds may lack the resources, time or knowledge necessary to facilitate these activities (e.g., Duncan, Kalil, Mogstad & Rege (2023)).[1] Parenting programs tailored for disadvantaged families have also proven effective in fostering school readiness (Attanasio et al. 2022, Duncan et al. 2023).[2]

This paper explores whether effective schools can serve as equalizers for children exposed to diverse rearing practices at home. A challenge in addressing this question lies in the fact that children are not randomly assigned neither to rearing nor to teaching practices, making it hard to interpret causally the evidence from observational studies. To circumvent this issue, we draw upon recent literatures demonstrating how to effectively engage parents in their children's learning through interventions that use text messages (e.g., Mayer, Kalil, Oreopoulos & Gallegos (2019)), and how to effectively support teachers to foster learning in school (e.g., Popova, Evans, Breeding & Arancibia (2022)).

Building on these literatures, we present evidence from a sequence of interventions, where we randomly offered parents and teachers of disadvantaged children in Colombia the possibility to adopt these effective rearing and teaching practices. Specifically, we randomly allocated

[1] High-quality preschool education can play a pivotal role in fostering learning and bridging these gaps for disadvantaged children (Heckman et al. 2010).

[2] In a standard parenting program, trained staff collaborates with families on activities encompassing feeding and care practices, child-rearing, and the psychosocial stimulation of their children. The effectiveness of these programs hinges on the curriculum, the regularity of meetings, and the active participation of parents.

children to a sequence of treatment assignments: only the parenting intervention (for the 3-month summer break before first grade), only the teacher intervention (first and second grades), both interventions, or no intervention at all. We then study the effects of these treatments on children's learning trajectories in primary school.

The first intervention - henceforth the parenting intervention - focused on fostering parental engagement with their children and improving their children's school readiness prior to entering first grade, specifically foundational numeracy skills. Over a three-month period prior to the beginning of the first grade, parents of children randomized to the parenting intervention were encouraged via phone-delivered messages to undertake several developmental-appropriate activities, that could support their children's learning and development. These activities tapped into daily family routines and involved the use of household objects, such as plates or cups (Mayer et al. 2019, York et al. 2019, Doss et al. 2019).

The second intervention - henceforth the teacher intervention - provided professional development opportunities to primary school teachers. Its objective was to improve teachers' pedagogy and foster children's math learning (Popova et al. 2022). Our intervention featured several in-person training sessions for teachers (six workshops of seven hours each distributed throughout year), as well as weekly in-person coaching in the classrooms, and online support throughout the school year.[3]

The two interventions aimed to foster early numeracy skills by engaging children in spontaneously developing mathematical competencies through structured manipulative activities, hence bridging the gap between symbolic and non-symbolic math (Dillon, Kannan, Dean, Spelke & Duflo 2017).[4] For example, care-givers were trained to first introduce sequences using

[3]Within treatment schools, we trained both first- and second-grade teachers. Second-grade teachers were trained in the school year in which children in our experiment entered second grade. So that children randomized to the professional development program were exposed to trained teachers over two consecutive school years.

[4]*Symbolic* mathematics allows humans to represent and describe the logic of the world around us with numbers and symbols. This ability to approximately compare, estimate, and manipulate large non-symbolic numerical quantities without language or symbols is known as *non-symbolic* math. Although we typically think about math symbolically, humans spontaneously exhibit numerical competences without reference to formal mathematics (Dehaene 2011). The two interventions were designed to develop children's formal numerical competences via games and manipulative activities, building on their intuitive math. For example, teachers were trained to first introduce sequences using every-day objects and colors to lay into the non-symbolic math and then make the connection to the sequences expressed with numbers. Similarly, parents were encouraged to introduce numbers by counting objects commonly available in every home, such as plates or cutlery.

every-day objects and colors to lay into the non-symbolic math and then make the connection to the sequences expressed with symbolic numbers (e.g., 2, 4, 6 etc.).

To study whether the interventions encouraged the behavioral changes we intended, we collected detailed information on parents' and teachers' behaviors using phone surveys with parents and in-person interviews with teachers as well as direct classroom observations of teachers and students. To measure the effects of the programs on children's learning, we measured learning outcomes in mathematics and literacy at end of first grade, and beginning and end of second grade using standard achievement tests.

During the summer break, parents of children participating in the parenting intervention engaged more actively with educational materials, read, told stories, and counted more than parents of children in the control group. Similarly, teachers in treatment schools changed their pedagogy in ways consistent with the program philosophy. They had more participative classroom sessions, changed appropriately their perception of the learning curriculum, used more manipulatives and play-based learning activities in their classrooms, and engaged more with the relevant academic material.

Studying children's standardized test scores, we find that children who received only the parenting intervention improved their mathematics learning by 0.18 of a standard deviation compared to children that received no interventions (p-value$<$0.0001). At the same time, children allocated only to the teacher intervention improved their mathematics learning by 0.22 of a standard deviation (p-value$<$0.0001).

The effects of the two interventions are both similar in magnitude, and not statistically different from each other (p-value=0.647), suggesting that effective classroom teaching can compensate for low levels of parental engagement at home. We also find that treatment effects for children that receive both interventions were similar in magnitude to, and non-statistically different from, the effects on children assigned to the teacher intervention only, suggesting that the two interventions acts as partial *substitutes* rather than *complements*.

The cost-benefit analysis reveals that the teacher professional development can be more cost-effective than the parental engagement. Namely, the parental engagement fosters chil-

dren's numeracy on average by 0.90 S.D. for every 100 USD per student, while the teacher development on average by 0.43-3.90 S.D. for every 100 USD per student (depending on the depreciation rate and teachers' turnover).[5] Importantly, teachers can reach the whole classroom and can teach more than one cohort of students. This finding calls for training teachers and improving their quality, as such interventions have the potential to serve as equalizers for children exposed to diverse rearing practices at home.

This paper contributes to the literature by bridging two strands of work on effective parenting practices and effective classroom teaching. Specifically, we exploit a sequence of interventions, that are randomized, to examine if effective schools can serve as equalizers for children exposed to diverse rearing practices. This, in turn, connects to the literature, underscoring the significance of early childhood in the long-term development of human capital (Carneiro & Heckman 2002, 2003, Berlinski & Vera-Hernández 2019). We elaborate below on how each intervention contributes on its on own to the two strands of the literature we bridge.

Our first intervention contributes to the literature, showing that text messages are effective to boost learning by engaging parents and boosting parent-child interactions (York et al. 2019, Doss et al. 2019, Mayer et al. 2019, Bergman 2019, Barrera et al. 2020, Bergman & Chan 2021, Cortes et al. 2021, Berlinski et al. 2022, Hernandez-Agramonte et al. 2023).[6] Relatively to these studies, we collect long-term data on standardized test scores and document persistent long-term effects of these intervention (as opposed to the impact immediately after the end of the intervention).

Our second intervention builds upon the literature, showing that high-quality teachers have a major impact on students' performance, both over the course of the school year (Araujo et al. 2016) and into adulthood (Chetty et al. 2014). Teacher professional developments have been shown to be an effective tool to boost learning in developed and developing countries (Popova et al. 2022, Alvarez Marinelli et al. 2023, Cilliers et al. 2020, Grönqvist et al. 2021, Rege et al. 2024, Di Tommaso et al. 2024). We contribute to these studies by providing training for teachers

[5] According to a study conducted by Bando et al. (2019), 10 interventions implemented in Latin America and the Caribbean improved math test scores by 0.55 S.D. for every 100 USD.

[6] Lynch et al. (2023) conducted a meta-analysis of 27 experimental and quasi-experimental studies on the effectiveness of summer programs as a means of improving mathematical skills in school-age children.

in subsequent grades, unlike other papers in the literature. This provides a more realistic picture of the potential impacts of teacher training on the whole profile of children's learning during primary school.

Finally, this paper connects to the literature on complementarities in learning by exploiting a sequence of interventions likely to affect children's learning (Cunha & Heckman 2007, Kinsler 2016, Johnson & Jackson 2019, Rossin-Slater & Wüst 2020, Carneiro et al. 2022, Meghir et al. 2023, Goff et al. 2023). We focus on a particular window of opportunity, the transition from kindergarten to first grade. The randomized control trial enables us to provide evidence of substitution rather than complementarities in math learning in this specific window of opportunity. We stress that the idea of dynamic complementarity involves computing the returns to different of investments over a broader set of windows of opportunities during the lifecycle. We therefore emphasize the importance of considering the multi-dimensionality of skills and investments as some dimensions could show complementarity, while others show substitutions.

The paper is organized as follows. Sections 2 and 3 present respectively the context and the experimental design, while Section 4 presents the results. Section 5 concludes.

2 Fostering children's numeracy via parents and teachers

2.1 Conceptual framework

Human capital formation has been shown to be important for many life course outcomes (Currie & Thomas 2001, Attanasio et al. 2022). Parents and teachers make choices about the environmental inputs to promote human capital formation subject to their constraints. While making these choices, they may face barriers, such as lack of resources, time, information, to create effective learning opportunities at home and school, with adverse implications for their children's human capital development in the long run. Importantly, before primary school enrollment, families are the main source of care and interaction for young children in the years(Currie & Almond 2011, Attanasio et al. 2022). In this context, formal schooling has the potential to level the playing field.

This paper studies how to support the development of children's numeracy skills. We focus on numeracy skills because these are important predictors of later academic achievement (Duncan et al. 2007) and fundamental to economic success in adulthood (Hanushek et al. 2015). Building early mathematical competence in pre-primary and primary education is therefore a key milestone of childhood development. This is particularly important in developing countries, where it is estimated that one in two children do not achieve basic numeracy competencies at the end of elementary school (UNESCO 2017).

Understanding how to alleviate such constraints, and specifically if (and how) these constraints interact, is important from both a theoretical and policy perspective to foster human capital when the returns are higher (e.g., Duncan, Kalil, Mogstad & Rege (2023)). To support children's learning and test if effective primary education can remediate early disadvantage, we implemented two interventions with children and their caregivers between 2020 and 2023: (i) a parenting intervention and (ii) a teacher professional development program.[7] We discuss how each intervention tackles the constraints caregivers can face in the next Sections 2.3 and 2.4. Section 3 presents the experimental design to understand if (and how) these constraints interact by exploiting the random assignment of children to a sequence of treatment combinations (i.e., only a parenting intervention, only a teacher intervention, both interventions, or no intervention).

2.2 Context

In Colombia, the context of this study, 53 percent of elementary school children scored at or below the minimal level of mathematical competency in 2017 according to the results of the national standardized test. Results in the PISA test for 15-year-old children the same year confirmed these low performance results (Schleicher 2018).

Our study took place in the municipality of Manizales, in the Department of Caldas in Colombia. Approximately 13.8 percent of Manizales residents live below the poverty line, and

[7]These interventions were implemented in partnership with the Academia Colombiana de Ciencias Exactas, Fisicas y Naturales of Colombia, Fundacion Luker, and PREST Math.

6.9 percent live in rural areas. As described in greater details below, roughly 70 percent of children in our sample can be considered socioeconomically disadvantaged.

2.3 The Parenting intervention

To improve children's school readiness, in the summer preceding their entry into first grade, we implemented *Aprender en Casa* (Learning at Home), a parenting intervention that combines culturally-appropriate learning materials and activities with online support - in the form of text messages - for children's primary caregivers.

We printed and delivered to every family that accepted to participate in our study (both treated and control households) a play book that included a set of well-structured activities that the parents could engage in with their child. The activities build on decades of research in education, and were designed by a team of international experts from PREST Math in collaboration with the Colombian Academy of Sciences with the objective of enriching the care and stimulation that parents provided to their children at home.

The activities tapped into daily family's routines and involved the use of commonly available household objects, such as plates, cups or cutlery. In the design of the book and the activities, special emphasis was given to building children's foundational numeracy skills through play. The activities in the booklet also covered the development of early literacy skills, time management (i.e., building routines for the child), fine and gross motor skills. For example, some activities invited children to color figures by drawing between lines to develop fine motor skills.[8]

We distributed the learning materials to all families in our study areas for ethical and logistical reasons, and encouraged a set of randomly selected households to engage in the activities proposed in the book via a randomized control trial. The set of randomly-selected households received weekly text messages to encourage its use, building on recent evidence showing that parents can be effectively supported in their role of caregivers using text messages (Mayer et al.

[8]Developing fine motor skills, such as being able to hold a pen, can help children during the process of learning when they enter first grade.

2019, York et al. 2019, Doss et al. 2019, Angrist et al. 2020). The messages were designed to enrich parent-child interactions by complementing the learning activities in the booklet.

The design of the text-messaging-intervention follows closely the work by York et al. (2019) in that treated parents received three messages per week. However, the content of the texts was adapted to the learning activities included in the booklet and to the local culture and context. Over the twelve-week summer break period, treated parents received a message on Mondays that describes a *fact* related to children's learning and development. This message was designed to provide information to parents about the importance of building specific skills for children's success in school. On Wednesdays, parents received a *tip* message, describing how to implement games and activities related to the *fact*. The activities described in these messages were designed to tap into the household routines and easily performed using household objects or the intervention booklet. Finally, on Fridays, parents received a *growth* message that reinforces the first two messages with some encouragement.[9]

The text messages were designed to support parents by reducing informational and behavioral hurdles that may constrain their engagement with their children. On the one hand, parents might have limited information about what skills their children are expected to develop before entering first grade. The *fact* message aimed at addressing this constraint. On the other hand, engaged parenting could be complicated or cognitively demanding, and families could be unsure about how to best support their children through developmentally appropriate activities. The *tip* and *growth* messages meant to ease these behavioral constraints.

2.4 The Teacher intervention

To support children's learning in school, we implemented a professional development program for first- and second-grade teachers. The program was designed to provide teachers with ped-

[9]For example, we send the following three text messages in week 1: (i) FACT: Counting is one of the first math skills that children learn. You can help by counting items during everyday activities. (ii) TIP: As you put the dishes away, count the plates 1-by-1 with your child. Try it again with bowls. Can your child count the cups by himself (herself)? Forks? (iii) GROWTH: Keep counting. You're preparing your child! You can count many things together - your fingers & toes or the number of steps to the car.

agogical tools to specifically support the development of children's numeracy competencies.[10] Our intervention featured several in-person training sessions for teachers (six workshops of seven hours each distributed during the school year), as well as weekly in-person coaching in the classrooms, and online support throughout the school year, following the best practices in the literature (Popova et al. 2022). Specifically, the training sessions introduced teachers to the structured activities, that they could replicate in class with their students.

Several distinct features characterize the pedagogical approach upon which the program builds. First, in contrast with more traditional learning techniques, that rely on rote learning and memorization of facts, the curriculum encourages an active pedagogy, in which children are stimulated to co-construct knowledge though exploration and interactions with their teachers and peers. Second, in line with this active pedagogy, the curriculum uses a set of materials and manipulative activities that facilitate children's active exploration of the environment.

These features of the curriculum are designed to help children learn mathematics by levering their intuitive ability to approximately compare, estimate, and manipulate large non-symbolic numerical quantities without language or symbols (Dehaene 2011). For example, teachers were trained to first introduce sequences using every-day objects and colors to lay into the non-symbolic math and then make the connection to the sequences expressed with symbolic numbers (e.g., 2, 4, 6 etc.).

To facilitate the activities involved in the curriculum, teachers also receive a work guide and a set of materials for concrete and semi-concrete manipulative activities. The same materials were also distributed in control schools, but without formal training for teachers on how to use them. The fact that classroom materials were made available to both treated and control schools ensures that we can isolate the impact of teacher professional development from the impact of receiving additional classroom resources.

We trained first-grade teacher throughout the first year of the program. In the following year, when children moved to the second grade, we also trained second-grade teachers. Therefore, children in treatment schools were exposed to trained teacher over two consecutive school years.

[10]The intervention was delivered by our project partner Fundacion Luker, a Colombian NGO committed to improve educational quality in the region, which ran and managed the delivery of the program activities.

3 Experimental design & methods

3.1 Experimental design

We conducted a randomized control trial with children randomly assigned to a sequence of treatment combinations: only a parenting intervention, only a teacher intervention, both interventions, or no intervention. An overview of the design is described below:

- We invited primary schools principals to take part in the study in 2021. Principals received official letters from the Secretary of Education that invited them to participate. 50 schools consented to participate. We randomly assigned half of the schools to receive the teacher intervention in the first grade in 2022 and in the second grade in 2023 (treatment group) and the other half received the teacher intervention in the first grade in 2023 and in the second grade in 2024 (control group). Randomization was stratified by school size (first stage school-level randomization). Following the recommendations in Athey & Imbens (2017), we included at least two treated and two control units in each strata.[11]

- Primary education institutions in Colombia include a compulsory kindergarten year of schooling called *Transicion* that caters children aged 5 to 6 years (OECD 2016). We identified all children enrolled in kindergarten classes in public schools using official administrative records. These administrative records also included basic socio-demographic characteristics of children and their households, such as child gender, age, and household socio-economic status. Using these records, we randomly assigned children to the parenting intervention using the same stratifying variables we used to randomize schools (second stage individual-level randomization) as well as stratifying by child's gender and enrolment batch (i.e., early vs late enrolment). This process resulted in 918 children assigned to the treatment group that received the parenting intervention and 893 children assigned to the control group that did not, for a total of 1811 children.[12]

[11] One strata had only two schools.

[12] After the randomization, we realized 13 of these 1811 preschool children were duplicates in the administrative records and deleted these observations from the analysis.

- The cross randomization of the two treatments resulted in 440 children being randomly assigned to receive both interventions, 478 children receiving only the parenting intervention, 452 children receiving only the teacher intervention, and 441 children receiving no intervention.

Parents were not informed about the treatment status of their school. Neither first or second grade teachers were informed if a child received the preschool booklet and the messaging treatment assignment.

We note that, by design, control schools include both children randomized *in* and *out* of the parenting intervention. If there are positive spillovers from treated to control children in the parenting intervention, we would only identify a lower bound on the effect of the parenting intervention on children's outcomes. In Section 4.2.3, we use the fact that our randomization generated exogenous variation in the proportion of children exposed to the parenting intervention within schools and show that there are no spillover effects from the program.

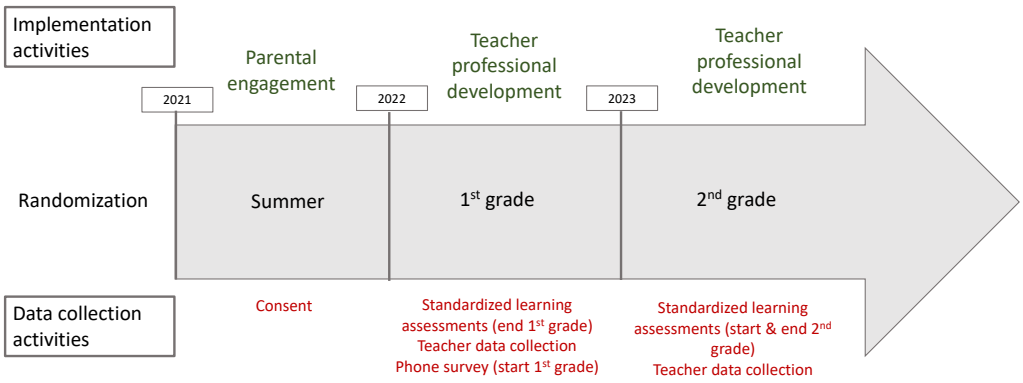
3.2 Timeline

The Colombian school year starts in February and ends in November. Figure 1 presents the timeline of our study, including data collection activities. We implemented the parenting intervention between December 2021 and February 2022, in the summer break preceding children's enrollment in first grade. The enrollment of first-grade teachers in the teacher professional development program started with an introductory day in November 2021, at the end of the school year. The first session of the teacher development was held in January 2022. Then, throughout the year, teachers received a total of five full-day in-person workshops (6 hours each) delivered by pedagogical experts. Teachers further received weekly in-class support by tutors during the school year, as well as online support. Second-grade teachers were trained in similar conditions and received similar support throughout 2023, when children were enrolled in second grade.

We collected data from parents, teachers and students. First, we run a phone-survey with parents after the parenting intervention in February 2022.Second, we interviewed teachers in schools and conducted direct classroom observations with teachers and children in October

2022 and 2023. Finally, we use data on standardized learning assessments conducted in schools in October 2022, February 2023 and October 2023. The survey instruments are described in detail in the next Section 3.3.

Figure 1: Timeline



3.3 Data

We collected rich data from parents and teachers to monitor engagement with the interventions (Das et al. 2013), as well as data on students' achievement in school.

3.3.1 Parent interviews

We collect data from parents in February 2022. Besides including basic socio-demographic information, the survey collected detailed information on parent-child interactions adapted from the HOME questionnaire (Bradley et al. 1988). Specifically, we collected information on the type and frequency of activities that the caregivers performed with their child in the seven days preceding the interview, such as counting numbers, singing, telling stories, or reading books. We also collected parental reported use of the learning materials that we distributed in November 2021. Interviews were conducted over the phone by trained enumerators, who were blind to the treatment status.

3.3.2 Teacher interviews and classroom observation

We collect data on teachers' practices at the end of first and second grade (i.e., around October 2022 and 2023) using two different instruments. First, we conducted a teacher interview, in which we asked teachers about curricular knowledge and classroom practices (e.g., classroom and behavioral management, learning expectations for their students, and pedagogical content knowledge).

Second, we coupled these self-reported measures with direct observation of teacher activities in the classroom, using the Stallings observation tool. Trained enumerators visited study schools and spent about one hour at the back of a classroom without disturbing the class and coded whether the teacher or the students were taking specific actions during the period of observation. During each visit, enumerators took different "snapshots" of what was happening in the class, and the extent to which students were actively engaging with the learning materials, whether they were distracted, and the type of resources and materials being used by teachers and students in the classroom.

3.3.3 Child survey

Learning assessments were administered in schools by trained enumerators, who were blind to students' treatment status and interviewed students individually. This is part of regular data collection data carried out on the universe of elementary school children by Fundacion Luker and the Secretary of Education of Manizales. These measures are used to evaluate the progress of education in the Municipality. In this sense, the outcomes of interest for this study were pre-specified. Other papers in this same setting have also used these same outcomes (see, for example, Barrera-Osorio et al. (2020) and Alvarez Marinelli et al. (2024)).

Assessments were administered orally at each school by trained enumerators using a tablet in a one-on-one setting. Most of the tasks in EGRA and EGMA are timed. Students are supposed to complete these tasks within one minute. The enumerator keeps track of whether the answer to each question is correct or incorrect. On average, each test takes about 20 minutes per student.

Since the intervention studied in this paper was a numeracy program, the EGMA was the natural target outcome. Students were assessed in October 2022 (at the end of first grade), February 2023 (at the beginning of second grade) and October 2023 (at the end of second grade). Children's numeracy and literacy skills are evaluated using the Early Grade Mathematics Assessment (EGMA) and Early Grade Reading Assessment (EGRA) (RTI-International 2009).

Outcome variables Using children's answers to the EGMA and EGRA assessments, we construct two main outcome variables, that measure respectively children's early numeracy and early literacy skills:

1. Numeracy index: we compute this index as the unstandardized sum of children's test scores on the number writing, number reading, number comparison, additions, subtractions, sequence, equations, and problem solving scales of EGMA. We standardized this index to have a mean of zero and standard deviation of one for children in the control group.

2. Literacy index: we compute this index as the unstandardized sum of children's test scores on the letter, sound, words, passage reading, and comprehension scales of EGRA. We standardized this index to have a mean of zero and standard deviation of one for children in the control group.

Below, we also report dis-aggregated treatment impacts on different sub-scales of EGRA and EGMA. In particular, we report treatment effects separately for basic skills that children are expected to master by the end of first grade (such as being able to recognize and write one and two digits numbers), and more more advanced skills (such as the ability to add and subtract one and two digits number, or identify which number completes a sequence). We further show that our results are robust constructing our main outcome variables using principal component analysis rather than simple unstandarized sum of correct answers.

3.4 Integrity of the experiment

3.4.1 Baseline balance

We assess the internal validity of our experiment by evaluating any differences in pre-treatment characteristics of parents, schools, teachers, and students. We find, that there are no noticeable unbalances. Appendix Table A1 shows balance in pre-treatment characteristics.

Table 1: Baseline Balance

	Control		T1		T2		T3		P-value					
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	C = T1	C= T2	C=T3	T1=T2	T1=T3	T2=T3
Panel A: Child characteristics														
Age	5.13	0.42	5.14	0.44	5.13	0.41	5.16	0.40	0.45	0.64	0.17	0.85	0.47	0.27
Disability	0.00	0.07	0.01	0.12	0.01	0.09	0.01	0.10	0.34	0.49	0.51	0.54	0.57	0.98
Low SES	0.68	0.47	0.65	0.48	0.75	0.43	0.71	0.45	0.29	0.48	0.91	0.29	0.61	0.17
Cognitive	0.57	0.26	0.58	0.25	0.62	0.26	0.62	0.26	0.67	0.11	0.08	0.18	0.15	0.98
Language	0.69	0.22	0.68	0.21	0.71	0.21	0.71	0.20	0.59	0.42	0.29	0.11	0.06	0.66
Motor	0.63	0.23	0.64	0.24	0.62	0.24	0.62	0.23	0.56	0.96	0.60	0.72	0.41	0.38
Socioemotional	0.85	0.20	0.84	0.18	0.86	0.17	0.86	0.16	0.94	0.37	0.58	0.35	0.57	0.44
Panel B: School characteristics														
Multigrade	0.04	0.19	0.04	0.20	0.03	0.17	0.03	0.18	0.74	0.17	0.22	0.15	0.19	0.15
Rural	0.07	0.26	0.07	0.26	0.15	0.36	0.15	0.36	0.17	0.24	0.21	0.19	0.16	0.62
Number of classes	2.66	1.87	2.69	1.86	2.10	0.89	2.07	0.90	0.20	0.36	0.37	0.33	0.33	0.45
Average class size	20.97	13.02	21.22	13.09	19.99	10.97	19.62	10.19	0.70	0.97	0.99	0.97	0.96	0.88

Notes: The table reports the mean and standard deviation of several observable characteristics across groups, together with the p-values for the equality of means across groups. The p-value is for a test of equality of means, after controlling randomization strata fixed effects.

3.4.2 Attrition and survey nonresponse

The experiment is not contaminated by attrition. No school left the sample during the first year and only two schools in the control group left in the second year.

Survey response rates in collection of students' outcomes were around 65% in the control group. Appendix Table A2 shows that there is no difference in survey nonresponse by treatment arm.

3.5 Empirical specifications

We estimate Intention-to-treat (ITT) effects on children's learning outcomes, using the following empirical specification:

$$y_{i,s,b,t} = \alpha + \beta_1 T_{1,i} + \beta_2 T_{2,s} + \beta_3 T_{3,i,s} + X'_{i,s,t}\gamma + \delta_b + \omega_t + \epsilon_{i,s,b,t} \quad (1)$$

where $y_{i,s,b,t}$ is the outcome of child i, in school s, strata b, measured at t, $X_{i,s,t}$ is a vector of child characteristics, δ_b and ω_t are strata and survey wave fixed effects.

$T_{1,i}$ is an indicator variable taking value of one for children randomized to receiving only the parenting intervention (i.e., children from control schools who were randomly assigned to the parenting program), $T_{2,s}$ is an indicator variable that takes value of one for children in schools randomized to receiving only the teacher intervention (i.e., children from treatment schools that did not receive the parenting intervention), and $T_{3,s,i}$ is an indicator variable that takes value of one for children in treatment schools that also received the parenting intervention. Our reference group (control group) therefore includes children in control schools, who were randomized to not receiving the parenting intervention. So treatment effects should be interpreted with respect to this group of children. We discuss bounds for potential spillover effects from the parenting intervention in Section 4.2.3. We cluster the standard errors at school level. When we report the estimates for the sub-indices of the numeracy and literacy indices, we report the Romano-Wolf adjusted p-value to account for multiple hypothesis testing.

β_1, β_2 and β_3, in equation (1) identify ITT effects, namely the effect of being randomly

assigned to a particular treatment combination. To benchmark the magnitude of the effects, we report the program impacts in terms of standard deviation (S.D.) units of the outcome variable of the control group over the observation window.

In our main specification, we pool children's data from the three different data collections together (end of first grade, beginning and end of second grade), but also report heterogeneity in treatment impacts over time below. We further study heterogeneity in treatment effects by child gender, and household socio-economic status (SES).

We estimate ITT effects of the teacher intervention on teachers' outcomes using the following specification:

$$y_{i,s,b,t} = \alpha + \beta_2 T_{2,s} + X'_{i,s,t}\gamma + \delta_b + \epsilon_{i,t} \tag{2}$$

where $y_{i,s,b,t}$ is the outcome of teacher i, in school s, strata b, measured at endline t, X_{it} is a vector of teacher characteristics, and $T_{2,s}$ is an indicator variable taking value of one if the teacher is a treatment school.

4 Results

This section presents our results. We first present the effects of the two programs on parent-child interactions and teachers' practices, to demonstrate that each of the two interventions was effective at improving household/classroom practices. We then study the impacts of the different treatment combinations to answer the question of whether effective teachers can serve as equalizers for children exposed to diverse rearing practices at home.

4.1 Caregivers' practices

4.1.1 Parents

First, Table 2 shows that the parental engagement treatment increases the index of parent-child interactions by 0.15 SD. While the increase is statistically significant at 5% level, we notice that a high proportion of parents in the control group report doing the activities that form the index

(Columns 2-6). We find a high self-reported level of engagement with the learning materials distributed among parents assigned to not receiving the parenting intervention: 77% of parents in this group reported that they used the play book at some point during the summer break. Parents in the treatment group were 5 percentage points more likely to report engaging with the materials (the difference between the two groups is significant at the 10% level).

Table 2: Intention-to-treat: parent-child interactions

	Index	Use booklet	Count	Read	Story	Sing
Parental engagement	0.146**	0.057*	0.052	0.069***	0.049	0.001
	(0.072)	(0.031)	(0.043)	(0.024)	(0.033)	(0.037)
Mean control	0.000	0.770	0.667	0.749	0.687	0.654
Observations	803	803	803	803	803	803

Note. The table presents the effects of the parenting intervention on parent-child interactions. We pool responses from the parent interview at the start and end of grade 1. Column 1 report treatment effects on a summary index of parent-child interactions constructed by combining the dummy variables for the parent-child interactions in columns 2-6. Controls include: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses *** p<0.01, ** p<0.05, * p<0.1.

4.1.2 Teachers

Table 3 reports treatment impacts of the teacher intervention on teacher's practices and behaviors. We find evidence that the program improved targeted teacher practices. Column 1 reports a 0.853 S.D. improvement on a summary index of teachers' behavior constructed from aggregating columns 2-7.

The remaining columns report treatment effects on different dimensions of teachers' behaviors (Appendix Table A3 reports the definition of each variable). Among teacher in control schools, 48% of them reported engaging with the learning materials that we distributed at some point during the school year. The proportion of teachers who report using the materials is almost double in treatment schools (94%), with the difference between the two groups being statistically significant at the 1% level (Column 2). Treated teachers deliver classes that are better structured, for example, by presenting the objectives of the class at the beginning and summarizing them at the end (Column 3). Treated teachers know the curriculum better (Column 4) and adopt pedagogical strategies that feature the active participation of students and the use of manipulative activities (Column 5). We also find suggestive evidence that children in

treatment schools are less likely to be distracted (Column 6). Finally, we find that teachers in treatment schools are also more likely to hold learning expectations for their students that are more aligned with the national curriculum (Column 7).

Appendix Tables A12 and A13 report the effect of (exogenous) share of students assigned to the parental engament on teacher practices in the full sample and the treated schools respectively.

Table 3: Intention-to-treat: teachers' practices (standardized)

	Index	Use teacher material	Class preparation delivery delivery	Curriculum knowledge	Manipulative activities	Distracted students	Learning expectation
Teacher development	0.853***	0.449***	0.170*	0.739***	0.345**	-0.139	0.518***
	(0.213)	(0.105)	(0.085)	(0.234)	(0.132)	(0.143)	(0.185)
Unstandardized mean control	0.000	0.475	0.000	0.000	0.000	0.000	0.000
Observations	123	123	123	123	123	123	123

Note. Column 1 is a summary index of teacher-child interactions constructed by combining the dummy variables for the teacher practices in columns 2-7. Column 2 is the use of the teacher development material. Column 3 is class preparation/delivery (index constructed from enumerator's observations on how class is delivered). Column 4 is curriculum knowledge (index constructed from responses on whether teachers know what students should learn). Column 5 is knowledge of manipulative activities (index constructed from responses on whether teachers use manipulative activities). Column 6 is distracted students (index constructed from responses on how often students are distracted). Column 7 is learning expectations regarding students and math (index constructed from responses on what students are expected to learn). Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Intention-to-treat: teachers' practices

	Index	Active class score	Class preparation delivery	Students' participation	Class disruption	Use teaching material	Learning expectations	Curriculum content knowledge
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment school	1.067***	0.446***	0.248*	0.143	-0.422**	0.514***	0.513***	0.803***
	(0.182)	(0.110)	(0.130)	(0.182)	(0.194)	(0.088)	(0.145)	(0.204)
Observations	119	713	117	117	117	116	117	117
Control group mean	-0.000	-0.000	-0.000	-0.000	-0.000	0.444	-0.000	0.000

Notes: The table reports the ITT effects of the teacher intervention on teacher's practices. Each column is a different regression for the effect on the teacher program on the outcome in the column header. Column 1 is a summary index of teacher's behavior (based on the outcome variables in columns 2 to 8). The outcome in column 2 is constructed using the Stallings observation tool and capture active teaching and learning during the class (e.g., students' asked questions or explained their answers etc.). The outcome in column 3 measures the quality of the class delivered by the teacher based on an assessment of the enumerator. The outcome in column 4 is an index of students' active participation in class. Column 5 measures disruption in the classroom. The outcome in column 6 is an indicator for whether the teacher uses the teaching materials distributed at the beginning of the school year in their classes. The outcome in column 7 measures teachers' learning expectations for their students (e.g., whether the students will do well in mathematics). The outcome in column 8 measures curriculum content knowledge. Column 2 to 5 are based on enumerator's direct observations. Columns 6 to 8 are based on teachers' report. All the outcome variables (except that in column 6) have been standardized to have a mean of zero and standard deviation of one in control schools. We pool data from grade 1 and grade 2 together, except for the Stallings observation tool that was only used in grade 2. The number of observations in column 2 is larger because this regression is based on the Stallings tool that features 10 observations per teacher. Standard errors clustered at the school lever are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

4.2 Students' outcomes

4.2.1 Test scores

Table 5 reports the treatment effect on children's numeracy skills. Column 1 presents the result for the summary index, while Columns 2-4 report treatment effects by distinguishing between basic skills that children are expected to master by the end of first grade (such as being able to recognize and write one and two digits numbers, Column 2), and more more advanced skills (such as the ability to add and subtract one and two digits number, Column 3, and identify which number completes a sequence, solving simple equations, and problem solving, Column 4).

Column 1 shows that, each of the two interventions in isolation improves children's numeracy skills compared to the control group. The point estimates imply that the numeracy scores of children, who received only the parenting intervention are 0.181 S.D. higher compared to children in the control group (p-value$<$ 0.0001). Similarly, children exposed only to trained teachers, who did not receive the parenting intervention, improved their test scores by 0.217 SD compared to children from the control group (p-value$<$ 0.0001). The point estimates are both similar in magnitude and not statistically different from each other (p-value for the difference is 0.647).

Table 5: Intention-to-treat: Children's numeracy skills

	Index	**Number write, read, compare**	**Add & subtract**	**Sequence, equations, Problems**
Parental engagement	0.181***	0.138***	0.137***	0.206***
	(0.046)	(0.044)	(0.046)	(0.042)
		[0.017]	[0.017]	[0.002]
Teacher development	0.217***	0.128**	0.161**	0.328***
	(0.073)	(0.062)	(0.075)	(0.064)
		[0.080]	[0.080]	[0.001]
Both	0.137**	0.053	0.126*	0.210***
	(0.067)	(0.059)	(0.064)	(0.066)
		[0.283]	[0.080]	[0.017]
Observations	3471	3471	3471	3471
R^2	0.306	0.366	0.215	0.170
P-val T1=T2	0.647	0.872	0.768	0.090
P-val T1=T3	0.534	0.136	0.878	0.949
P-val T2=T3	0.105	0.019	0.505	0.159
Unstandardized mean control	0.436	0.723	0.252	0.289

Note. The table presents the effects of different treatment combinations on children's numeracy skills. The outcome variables have been standardized to have mean 0 and standard deviation 1 in the control group. Controls include: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Below each coefficient in square brackets we report the Romano-Wolf adjusted p-value to account for multiple hypothesis testing. P-values at the bottom of the Table tests that the equality of coefficients (parent intervention (T1), teacher intervention (T2) and both interventions (T3)). Standard errors clustered at the school level in parentheses. *** p$<$0.01, ** p$<$0.05, * p$<$0.1.

We further find that children in treatment schools who were randomized to the parenting intervention improved by 0.137 S.D. compared to children from the control group (p-value=0.045). We find limited evidence that the effect of receiving only one intervention is different from that of receiving both interventions, and - if anything - we find that the two programs are *substitutes* rather than complements. The p-values for the difference between receiving only one or both interventions are reported at the bottom of the table, and suggest that the impact of the combined interventions is no different from that of the parenting intervention alone (p-value=0.534), and 0.036 S.D. smaller (0.217-0.181) that that of the teacher intervention only (p-value=0.105)

Unpacking the treatment impacts on the numeracy index in its different components, we find smaller effects on basic numeracy skills and larger effects on more advanced skills. When considering children's mastery of simple skills, such as reading and comparing numbers, the parenting intervention improves children's outcomes by 0.138 S.D. and the teacher intervention by 0.128 S.D. The largest treatment impacts are found on the most advanced skills, that include children's ability to complete sequences, solve simple equations and problem solving. On this dimension, the parenting intervention improve children's competencies by 0.206 S.D., the teacher intervention only by 0.328 S.D., and both interventions by 0.210 S.D.

Table 6 presents the results for children's literacy skills. We find some evidence that parental engagement and teacher development also improve children's literacy scores. Column 1 suggests that each of the two interventions in isolation improves children's literacy skills compared to the control group, with such estimates not being statistically different from receiving both interventions. This improvement appears to be driven by parents and teachers improving children's ability to read and understand a passage (Column 4). It is not surprising that more parent/children engagement and better classroom management can affect literacy as we observe the largest improvement in the ability to solve equations, sequences and problems formulated with words.

We perform several robustness checks to our results. In Appendix Tables A4 and A5, we reproduce the main estimates in Tables 5 and 6 using factor scores instead of raw scores, and find

Table 6: Intention-to-treat: Children's literacy skills

	Index	**Letter initial sound**	**Simple pseudo words**	**Passage reading & comprehension**
Parental engagement	0.151*	0.097	0.129*	0.132*
	(0.077)	(0.066)	(0.069)	(0.068)
		[0.258]	[0.163]	[0.159]
Teacher development	0.149*	0.028	0.120	0.175**
	(0.087)	(0.080)	(0.074)	(0.081)
		[0.709]	[0.221]	[0.115]
Both	0.093	0.034	0.040	0.123
	(0.093)	(0.074)	(0.081)	(0.084)
		[0.709]	[0.709]	[0.258]
Observations	3471	3471	3471	3471
R^2	0.136	0.053	0.169	0.271
P-val T1=T2	0.980	0.419	0.914	0.603
P-val T1=T3	0.550	0.418	0.307	0.916
P-val T2=T3	0.368	0.892	0.147	0.419
Unstandardized mean control	0.356	0.455	0.291	0.263

Note. The table presents the effects of different treatment combinations on children's literacy skills. The outcome variables have been standardized to have mean 0 and standard deviation 1 in the control group. Controls include: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Below each coefficient in square brackets we report the Romano-Wolf stepdown adjusted p-value to account for multiple hypothesis testing. P-values at the bottom of the Table tests that the equality of coefficients (parent intervention (T1), teacher intervention (T2) and both interventions (T3)). Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

very similar results (if anything the point estimates are larger when we consider factor scores). In Appendix Table A6, we test the significance of our results using randomization inference, and find similar results. We also test for contamination bias as discussed in Goldsmith-Pinkham et al. (forthcoming) and do not find evidence for contamination bias for any of the treatment arms (Appendix Table A7).

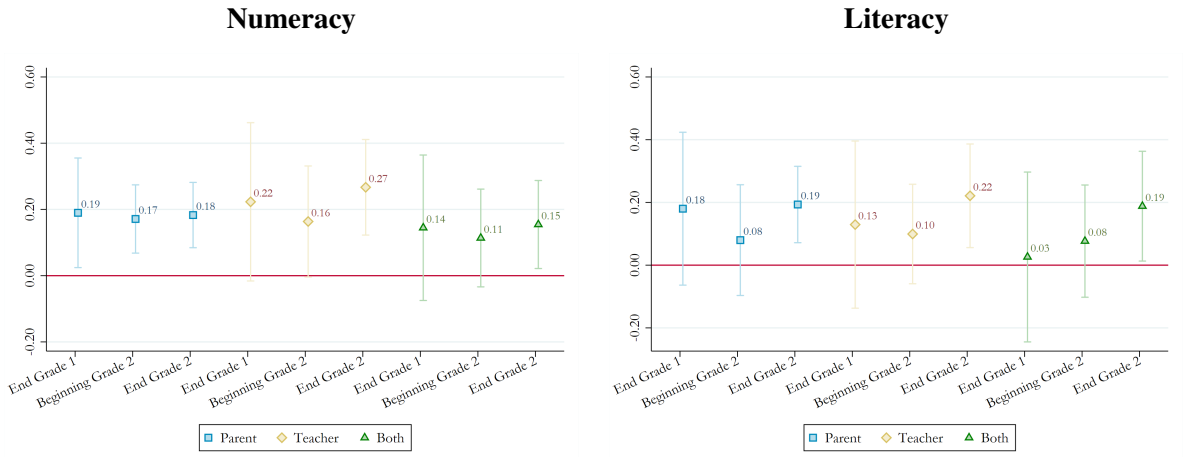
4.2.2 Heterogeneity

We investigate heterogeneity in treatment impacts along three margins. First, we consider if the treatment effects are heterogeneous over time across survey waves. Then, we consider if the effects are heterogeneous by child gender or socio-economic status (SES).

Figure 2 presents the main estimates from Column 1 in Tables 5 and 6 by school grade to examine if the effects fade in or fade out over time. We observe that the effects are stable over time, and if anything the teacher development seems to increase in the third wave of data collection.

Appendix Figure A1 presents the treatment effects on numeracy and literacy by gender. The effects are similar by gender. On the other hand, de Barros et al. (2024) and Di Tommaso et al. (2024) find effects of activity-based instruction only for girls in India and Italy. Appendix

Figure 2: Intention-to-treat estimates by wave: Numeracy and literacy skills (scores)



Note. The Figure presents the main estimates for numeracy and literacy skills for each treatment arm by wave. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level. Confidence interval at 95% level.

Figure A2 presents the treatment effects on numeracy and literacy by parents' SES. The parental engagement seems to be more effective for high SES, while the teacher development for low SES. These difference are not statistically different though.

Finally, Appendix Figure A3 presents the quantile treatment effects. We do not observe any large differences in treatment effects by quantiles.

4.2.3 Spillovers

As described earlier, given the relatively small number of schools that participated in the experiment, we do not have a pure control group, as children who did not received the parenting intervention in control schools could be indirectly affected by the parenting program through spillover effects. We exploit the fact that our second stage randomization generated exogenous variation in the proportion of children exposed to the parenting intervention within control schools to shed light on the existence of spillover effects from the program.

In Appendix Table A10, we limit the sample to the group of control schools, and regress the test scores of children randomized out of the parenting intervention on the (exogenous) proportion of children who received the parenting intervention in the school. We do not find evidence of spillovers.

4.3 Cost-benefit analysis

This section outlines the cost-benefit analysis of the parental engagement and teacher professional development.

The costs of the teacher professional development program fall into three categories: (i) program administration and staff, (ii) program materials and (iii) training. For all three categories, rates are available for 2022, 2023 and 2024. For 2025 onwards, rates are estimated using the Colombian inflation rates. The teacher professional development is set to run during five years for first- and second-grade teachers. Each grade has an implementation year and four additional continuation years. During the continuation years, there are recurring student and tutor costs, while teacher costs account for teacher turnover.

First, program administration and staff costs include tutor salaries, tutor training, tutor materials, and program administration. Second, material costs include all costs related to producing, translating, printing, and distributing teachers' and students' booklets as well as costs from purchasing and distributing support materials for both students and teachers. The production and translation costs are only incurred during implementation years. Teachers receive only one booklet and set of materials every five years, but we assume a teacher turnover of 20%. Every year all students receive a booklet and materials. Third, training costs include the cost of paying a replacement teacher during the workshops, plus the cost of each workshop per teacher. Teachers have six workshops during the first year. Afterwards, only new teachers receive the workshop.

The costs of the parental engagement fall into two categories: (i) program administration and staff and (ii) implementation and program materials. The costs are estimated only for the base year. Program and staff administration costs include the salary of a tutor available to provide support to parents during the program, that lasts 3 months, as well as a coordinator, hired for three months to oversee the implementation of the program. Implementation and program material costs include the production and translation of the booklet, printing and distributing it as well as sending 54 text messages per participant. We also consider the cost of a school kit needed to complete the activities of the workbook.

We find that the teacher professional development is more cost-effective than the parental engagement. The average cost per student is 21.30 USD for the teacher professional development and 20.03 USD for the parental engagement. On the other hand, we compute the benefits by taking into account that teachers can teach at least one cohort of students and training can depreciates for each new cohort. We assume a discount rate of 10%. The parental engagement fosters children's numeracy on average by 0.90 S.D. for every 100 USD per student. On the other hand, the teacher development can foster children's numeracy on average by 0.43-3.90 S.D. for every 100 USD per student, depending on the depreciation rate and teachers' turnover.

5 Conclusion

This paper experimentally studies whether effective schools can serve as equalizers for children exposed to diverse rearing practices. A challenge in addressing this question lies in the fact that neither rearing nor teaching practices are chosen at random, making it hard to interpret causally the evidence from observational studies. To circumvent this issue, we employ a randomized control trial and expose children to a sequence of intervention scenarios: only a parenting intervention, only a teacher intervention, both interventions, or no intervention. The parental engagement and teacher development interventions use manipulative activities to assist treated children in bridging the gap between non-symbolic mathematics - lacking numbers and symbols - and symbolic mathematics.

Our findings show that parental engagement and teacher development interventions act as substitutes, each resulting in an increase of approximately 0.20 standard deviations in math test scores measured through standardized assessments. We also present evidence of enhanced quality in parent-child interactions and improved teacher practices.

Parenting interventions play a crucial role in shaping children's outcomes during their formative years. Our research suggests that effective schools are already addressing many of the challenges that children may face when they enter the educational system. It is essential, therefore, to carefully explore and select parenting interventions that can complement the positive

work done by good schools. At the same time, our study serves as a reminder of the critical role effective schools play in bridging learning gaps, which can have broader implications for various aspects of life.

References

Almond, D., Currie, J. & Duque, V. (2018), 'Childhood circumstances and adult outcomes: Act II', *Journal of Economic Literature* **56**(4), 1360–1446.

Alvarez Marinelli, H., Berlinski, S. & Busso, M. (2024), 'Remedial education: evidence from a sequence of experiments in colombia', *Journal of Human Resources* **59**(1), 141–174.

Alvarez Marinelli, H., Berlinski, S., Busso, M. & Martinez Correa, J. (2023), 'Improving early literacy through teacher professional development: Experimental evidence from Colombia', *Journal of Public Economics Plus* **4**, 100019.

Angrist, N., Bergman, P., Brewster, C. & Matsheng, M. (2020), School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic, Technical report, National Bureau of Economic Research.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y. & Schady, N. (2016), 'Teacher quality and learning outcomes in kindergarten', *The Quarterly Journal of Economics* **131**(3), 1415–1453.

Athey, S. & Imbens, G. W. (2017), The econometrics of randomized experiments, *in* 'Handbook of economic field experiments', Vol. 1, Elsevier, pp. 73–140.

Attanasio, O., Cattan, S. & Meghir, C. (2022), 'Early childhood development, human capital, and poverty', *Annual Review of Economics* **14**, 853–892.

Bando, R., Näslund-Hadley, E. & Gertler, P. (2019), 'Effect of inquiry and problem based pedagogy on learning: Evidence from 10 field experiments in four countries', *National Bureau of Economic Research working paper* .

Barrera, O., Macours, K., Premand, P., Vakis, R. et al. (2020), *Texting parents about early child development: Behavioral changes and unintended social effects*, World Bank Group, Poverty and Equity Global Practice & Development Impact.

Barrera-Osorio, F., Gonzalez, K., Lagos, F. & Deming, D. J. (2020), 'Providing performance information in education: An experimental evaluation in colombia', *Journal of Public Economics* **186**, 104185.

Bergman, P. (2019), 'How behavioral science can empower parents to improve children's educational outcomes', *Behavioral Science & Policy* **5**(1), 53–67.

Bergman, P. & Chan, E. W. (2021), 'Leveraging parents through low-cost technology: The impact of high-frequency information on student achievement', *Journal of Human Resources* **56**(1), 125–158.

Berlinski, S., Busso, M., Dinkelman, T. & Martínez, C. (2022), 'Reducing parent-school information gaps and improving education outcomes: Evidence from high-frequency text messages', *Journal of Human Resources* pp. 1121–11992R2.

Berlinski, S. & Vera-Hernández, M. (2019), The rationale for interventions to foster child development, *in* 'Oxford Research Encyclopedia of Economics and Finance'.

Bradley, R. H., Caldwell, B. M., Rock, S. L., Hamrick, H. M. & Harris, P. (1988), 'Home observation for measurement of the environment: Development of a home inventory for use with families having children 6 to 10 years old', *Contemporary Educational Psychology* **13**(1), 58–71.

Carneiro, P., Cruz-Aguayo, Y., Pachon, R. H. & Schady, N. (2022), 'Dynamic complementarity in elementary schools: Experimental estimates from ecuador'.

Carneiro, P. & Heckman, J. (2003), 'Human capital policy', *IZA discussion paper* .

Carneiro, P. & Heckman, J. J. (2002), 'The evidence on credit constraints in post-secondary schooling', *The Economic Journal* **112**(482), 705–734.

Chetty, R., Friedman, J. N. & Rockoff, J. E. (2014), 'Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood', *American Economic Review* **104**(9), 2633–2679.

Cilliers, J., Fleisch, B., Prinsloo, C. & Taylor, S. (2020), 'How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching', *Journal of Human Resources* **55**(3), 926–962.

Cortes, K. E., Fricke, H., Loeb, S., Song, D. S. & York, B. N. (2021), 'Too little or too much? Actionable advice in an early-childhood text messaging experiment', *Education Finance and Policy* .

Cunha, F. & Heckman, J. (2007), 'The technology of skill formation', *American economic review* **97**(2), 31–47.

Currie, J. & Almond, D. (2011), Human capital development before age five, Vol. 4 of *Handbook of Labor Economics*, Elsevier, pp. 1315–1486.

Currie, J. & Thomas, D. (2001), Early test scores, school quality and ses: Longrun effects on wage and employment outcomes, *in* 'Worker wellbeing in a changing labor market', Emerald Group Publishing Limited, pp. 103–132.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K. & Sundararaman, V. (2013), 'School inputs, household substitution, and test scores', *American Economic Journal: Applied Economics* **5**(2), 29–57.

de Barros, A., Fajardo-Gonzalez, J., Glewwe, P. & Sankar, A. (2024), 'The limitations of activity-based instruction to improve the productivity of schooling', *The Economic Journal* **134**(659), 959–984.

Dehaene, S. (2011), *The number sense: How the mind creates mathematics*, OUP USA.

Di Tommaso, M. L., Contini, D., De Rosa, D., Ferrara, F., Piazzalunga, D. & Robutti, O. (2024), 'Tackling the gender gap in mathematics with active learning methodologies', *Economics of Education Review* **100**, 102538.

Dillon, M. R., Kannan, H., Dean, J. T., Spelke, E. S. & Duflo, E. (2017), 'Cognitive science in

the field: A preschool intervention durably enhances intuitive but not formal mathematics', *Science* **357**(6346), 47–55.

Doss, C., Fahle, E. M., Loeb, S. & York, B. N. (2019), 'More than just a nudge: Supporting kindergarten parents with differentiated and personalized text messages', *Journal of Human Resources* **54**(3), 567–603.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J. et al. (2007), 'School readiness and later achievement.', *Developmental psychology* **43**(6), 1428.

Duncan, G., Kalil, A., Mogstad, M. & Rege, M. (2023), 'Investing in early childhood development in preschool and at home', *Handbook of the Economics of Education* **6**, 1–91.

Goff, L., Malamud, O., Pop-Eleches, C. & Urquiola, M. (2023), 'Interactions between family and school environments: Access to abortion and selective schools', *Journal of Human Resources* .

Goldsmith-Pinkham, P., Hull, P. & Kolesár, M. (forthcoming), 'Contamination bias in linear regressions', *American Economic Review* .

Grönqvist, E., Öckert, B. & Rosenqvist, O. (2021), 'Does the "boost for mathematics" boost mathematics? A large-scale evaluation of the "lesson study" methodology on student performance', *Working Paper* .

Hanushek, E. A., Schwerdt, G., Wiederhold, S. & Woessmann, L. (2015), 'Returns to skills around the world: Evidence from PIAAC', *European Economic Review* **73**, 103–130.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A. & Yavitz, A. (2010), 'The rate of return to the Highscope Perry Preschool Program', *Journal of public Economics* **94**(1-2), 114–128.

Hernandez-Agramonte, J. M., Namen, O., Naslund-Hadley, E. & Biehl, M. L. (2023), 'Supporting early childhood development remotely: experimental evidence from SMS messages', *Journal of Development Economics* p. 103201.

Johnson, R. C. & Jackson, C. K. (2019), 'Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending', *American Economic Journal: Economic Policy* **11**(4), 310–349.

Kinsler, J. (2016), 'Teacher complementarities in test score production: Evidence from primary school', *Journal of Labor Economics* **34**(1), 29–61.

Lynch, K., An, L. & Mancenido, Z. (2023), 'The impact of summer programs on student mathematics achievement: A meta-analysis', *Review of Educational Research* **93**(2), 275–315.

Mann, H. (1848), 'Twelfth annual report to the Massachusetts Board of Education', *The republic and the school: Horace Mann and the education of free men* **123**.

Mayer, S. E., Kalil, A., Oreopoulos, P. & Gallegos, S. (2019), 'Using behavioral insights to increase parental engagement: The parents and children together intervention', *Journal of Human Resources* **54**(4), 900–925.

Meghir, C., Attanasio, O., Jervis, P., Day, M., Makkar, P., Behrman, J., Gupta, P., Pal, R., Phimister, A., Vernekar, N. et al. (2023), 'Early stimulation and enhanced preschool: A randomized trial', *Pediatrics* **151**(Supplement 2).

OECD (2016), 'Reviews of national policies for education: Education in Colombia', *OECD Publishing, Paris* .

Popova, A., Evans, D. K., Breeding, M. E. & Arancibia, V. (2022), 'Teacher professional development around the world: The gap between evidence and practice', *The World Bank Research Observer* **37**(1), 107–136.

Rege, M., Størksen, I., Solli, I. F., Kalil, A., McClelland, M. M., Ten Braak, D., Lenes, R., Lunde, S., Breive, S., Carlsen, M. et al. (2024), 'The effects of a structured curriculum on preschool effectiveness: A field experiment', *Journal of Human Resources* **59**(2), 576–603.

Rossin-Slater, M. & Wüst, M. (2020), 'What is the added value of preschool for poor children?

Long-term and intergenerational impacts and interactions with an infant health intervention', *American Economic Journal: Applied Economics* **12**(3), 255–286.

RTI-International (2009), 'Early grade reading assessment toolkit', *World Bank Working Paper, Office of Human Development* .

Schleicher, A. (2018), 'Insights and interpretations', *PISA 2018* **10**.

UNESCO (2017), 'More than one-half of children and adolescents are not learning worldwide', *UIS Fact Sheet* **46**.

York, B. N., Loeb, S. & Doss, C. (2019), 'One step at a time: The effects of an early literacy text-messaging program for parents of preschoolers', *Journal of Human Resources* **54**(3), 537–566.

Appendices to "Effective Families or Effective Schools? Experimental Evidence on Fostering Children's Numeracy"

A Appendix Tables and Figures

Table A1: Balance

Variable	Control mean (s.d)	OLS Coeff. (s.e.) - Parents	P-value - Parents	OLS Coeff. (s.e.) - Teachers	P-value - Teachers	OLS Coeff. (s.e.) - Both	P-value - Both	Obs.
Age (2021)	5.135	0.007	0.816	0.018	0.424	0.032	0.246	1757
	(0.411)	(0.028)		(0.022)		(0.027)		
Low strata (1=yes)	0.678	0.043	0.483	–0.024	0.289	0.007	0.908	1716
	(0.468)	(0.061)		(0.022)		(0.059)		
Disability (1=yes)	0.005	0.004	0.482	0.011	0.332	0.005	0.501	1721
	(0.069)	(0.006)		(0.011)		(0.007)		
Has Predictores 1 (1=yes)	0.924	–0.006	0.763	–0.022	0.321	–0.017	0.375	1796
	(0.265)	(0.020)		(0.022)		(0.019)		
Has 2021 SIMAT (1=yes)	0.970	–0.005	0.591	–0.007	0.519	–0.015	0.131	1796
	(0.170)	(0.010)		(0.011)		(0.010)		

Note. This table presents the balance estimates. Students are assigned to one of four groups: (i) only parental engagement (presented in columns 'Parents'), (ii) only teacher professional development (presented in columns 'Teachers'), (iii) both interventions, or (iv) no program (not presented). The control mean refers to the control group for all treatments. Predictores and SIMAT are administrative data from 2021. Predictores measures early development cognitive and socio-emotional skills. SIMAT has data on age and socio-economic strata. Column labeled 'Variable' describes the variable used to check balance. Column 'Control mean' presents the mean for the control group. Column 'OLS Coeff. (s.e)' shows the variable's coefficients from the OLS regression and its standard errors below in parenthesis. Fixed effects using the size of first grade in eligible schools, the student's gender, and if the student was new, that is, was added to our dataset when we collected the consent (stratification variables) are included. Column 'P-value' shows the p-value of a test of H0 : $\theta = 0$. Significance levels are indicated by: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A2: Attrition

Variable	Control mean (s.d)	OLS Coeff. (s.e.) - Parents	P-value - Parents	OLS Coeff. (s.e.) - Teachers	P-value - Teachers	OLS Coeff. (s.e.) - Both	P-value - Both	Obs.
Has EGRA EGMA W2 (1=yes)	0.673	0.050	0.113	–0.032	0.254	0.073**	0.011	1796
	(0.470)	(0.031)		(0.028)		(0.027)		
Has EGRA EGMA W3 (1=yes)	0.643	0.054	0.127	–0.044	0.198	0.024	0.427	1796
	(0.480)	(0.035)		(0.034)		(0.030)		
Has EGRA EGMA W4 (1=yes)	0.593	0.054	0.196	–0.035	0.153	0.034	0.415	1796
	(0.492)	(0.041)		(0.024)		(0.042)		
Has EGRA EGMA W5 (1=yes)	0.558	0.050	0.220	–0.011	0.701	0.047	0.304	1796
	(0.497)	(0.041)		(0.030)		(0.045)		

This table presents estimations for individual and combined treatments for all preschool students, regardless of whether their parents/guardians accepted the consent for the preschool program. Students are assigned to one of four groups: (i) only preschool treatment (presented in columns 'Parents'), (ii) only teacher training program (presented in columns 'Teachers'), (iii) both preschool and teacher training programs (presented in columns 'Both'), or (iv) no program (not presented). The control mean refers to the control group for all treatments. The 'W' in the EGRA EGMA variables stands for 'wave'. Wave 2 was collected in October 2022, Wave 3 in March 2023, Wave 4 in October 2023 and Wave 5 in March 2024. Column labeled 'Variable' describes the variable used to check balance. Column 'Control mean' presents the mean for the control group. Column 'OLS Coeff. (s.e)' shows the variable's coefficients from the OLS regression and its clustered standard errors at the school level (unit of randomization for teacher training treatment) below in parenthesis. Fixed effects using the size of first grade in eligible schools, the student's gender, and if the student was new, that is, was added to our dataset when we collected the consent (stratification variables) are included. Column 'P-value' shows the p-value of a test of H0 : $\theta = 0$. Significance levels are indicated by: * significant at 10%, ** significant at 5%, *** significant at 1%.

Table A3: Variable definition: Teacher practices

Variable name	Questionnaire items
Class preparation/delivery	Enumerator's observations on how class is delivered: (i) teacher presented the learning objectives at the beginning of the lesson, (ii) teacher gave instructions before every activity, (iii) teacher did a closure a the end of the class summarizing learning, (iv) material was prepared.
Curriculum knowledge	Responses on whether teachers know what students should learn: (i) only quantities of a maximum of one hundred, (ii) not memorize the addition and subtraction, (iii) vertically add and subtract, (iv) addition and subtraction of 3-digit quantities, (v) math expressions with missing number, (vi) problems formulated with words (vii) addition and subtraction plans.
Knowledge of manipulative activities	Agree that students should learn addiction and subtraction via: (i) using their fingers, (ii) working with concrete material.
Distracted students	Responses on how often students are distracted: (i) the teacher interrupts classroom to get students' attention, (ii) students are distracted by other activities, (iii) class starts late to get students' attention.
Learning expectations regarding students and math	Responses on what students are expected to learn: (i) all of your students can learn mathematics, (ii) students find it fun to learn mathematics, (iii) all of your students look forward to math class with (iv) math does not cause them distress for many students (v) some students have aptitude for mathematics, (vi) by the end of the year most of your students achieve the expected level in mathematics.

Table A4: Intention-to-treat: Numeracy skills (factors)

	Index	**Number write, read, compare**	**Add & subtract**	**Sequence, equations, Problems**
Parental engagement	0.193***	0.147**	0.146***	0.195***
	(0.051)	(0.055)	(0.046)	(0.039)
Teacher development	0.231***	0.147**	0.165**	0.332***
	(0.079)	(0.071)	(0.073)	(0.067)
Both	0.158**	0.061	0.150***	0.219***
	(0.072)	(0.069)	(0.053)	(0.073)
Observations	3471	3471	3471	3471
R^2	0.107	0.083	0.068	0.092
P-val T1=T2	0.632	0.999	0.807	0.066
P-val T1=T3	0.629	0.156	0.949	0.762
P-val T2=T3	0.235	0.021	0.824	0.225
Unstandardized mean control	0.436	0.723	0.252	0.289

Note. Indexes are computed with factor analysis on control group data. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Intention-to-treat: Literacy skills (factors)

	Index	**Letter initial sound**	**Simple pseudo words**	**Passage reading & comprehension**
Parental engagement	0.151*	0.098	0.148*	0.116
	(0.080)	(0.062)	(0.075)	(0.075)
Teacher development	0.161*	0.024	0.144*	0.178**
	(0.087)	(0.080)	(0.079)	(0.083)
Both	0.091	0.037	0.057	0.114
	(0.092)	(0.074)	(0.086)	(0.083)
Observations	3471	3471	3471	3471
R^2	0.071	0.038	0.070	0.061
P-val T1=T2	0.909	0.369	0.960	0.448
P-val T1=T3	0.519	0.408	0.295	0.979
P-val T2=T3	0.302	0.782	0.189	0.334
Unstandardized mean control	0.356	0.455	0.291	0.263

Note. Indexes are computed with factor analysis on control group data. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Intention-to-treat: Numeracy skills (randomization inference p-value)

	Index	**Number write, read, compare**	**Add & subtract**	**Sequence, equations, Problems**
Parental engagement	0.181***	0.138***	0.137***	0.206***
	(0.046)	(0.044)	(0.046)	(0.042)
	[0.000]	[0.003]	[0.002]	[0.000]
Teacher development	0.217***	0.128**	0.161**	0.328***
	(0.073)	(0.062)	(0.075)	(0.064)
	[0.000]	[0.000]	[0.000]	[0.000]
Both	0.137**	0.053	0.126*	0.210***
	(0.067)	(0.059)	(0.064)	(0.066)
	[0.001]	[0.133]	[0.005]	[0.000]
Observations	3471	3471	3471	3471
R^2	0.306	0.366	0.215	0.170
P-val T1=T2	0.647	0.872	0.768	0.090
P-val T1=T3	0.534	0.136	0.878	0.949
P-val T2=T3	0.105	0.019	0.505	0.159
Unstandardized mean control	0.436	0.723	0.252	0.289

Note. Indexes are computed with factor analysis on control group data. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Randomization inference p-values between square brackets.

Table A7: Contamination bias test

Panel A: full sample				
	OWN	ATE	EW	CW
Parental engagement	.	.	0.008	0.016
SE	.	.	0.013	0.027
Teacher development	-0.007	.	0.008	0.004
SE	0.032	.	0.020	0.022
Both	-0.019	0.020	-0.009	0.014
SE	0.033	0.026	0.020	0.019
Panel B: overlap sample				
	OWN	ATE	EW	CW
Parental engagement	0.003	0.005	0.008	0.023
SE	0.013	0.025	0.013	0.027
Teacher development	-0.008	0.008	0.010	0.012
SE	0.032	0.031	0.019	0.021
Both	-0.021	0.018	-0.008	0.018
SE	0.032	0.025	0.019	0.020

Note. Table presents the contamination bias test for the standardized numeracy score (Goldsmith-Pinkham et al. forthcoming). The table reports the difference between the PL estimator and the other estimators with respective standard errors in the full sample (panel A) and the overlap sample (panel B). A statistically significant difference between the the other estimators and PL estimate for each treatment (columns) in the full sample and in the overlap sample would mean significant contamination bias. PL is the partly linear model (PL). OWN is the own treatment effect component of the PL estimator that subtracts an estimate of the contamination bias. ATE is the unweighted average treatment effect, implemented using regression that includes interactions of covariates with the treatment indicators. EW is the Weighted ATE estimator based on easiest-to-estimate weighting (EW) scheme, implemented by running one-treatment-at-a-time regressions. CW is weighted ATE estimator using easiest-to-estimate common weighting (CW) scheme, implemented using weighted regression.

Table A8: Spillovers (only control children - no parental engagement): Numeracy skills

	Index	**Number write, read, compare**	**Add & subtract**	**Sequence, equations, problems**
Panel A: Spillovers within schools (full sample)				
% treated parental engagement (within school)	0.480	1.331**	-0.269	0.012
	(0.772)	(0.590)	(0.827)	(0.631)
Observations	1740	1740	1740	1740
R^2	0.298	0.358	0.218	0.166
Unstandardized mean control	0.507	0.834	0.300	0.324
Panel B: Spillovers within control schools				
% treated parental engagement (within school)	0.078	0.792	-0.566	0.022
	(1.005)	(0.755)	(1.090)	(0.581)
Observations	834	834	834	834
R^2	0.348	0.400	0.238	0.185
Unstandardized mean control	0.507	0.834	0.300	0.324
Panel C: Spillovers within treated schools				
% treated parental engagement (within school)	-0.195	0.185	-0.574	0.079
	(1.114)	(0.676)	(1.349)	(0.785)
Observations	906	906	906	906
R^2	0.324	0.392	0.252	0.214
Unstandardized mean control	0.540	0.848	0.323	0.331

Note. The table presents the estimates for possible spillovers in treated and control schools. We proceed in two steps: first, we compute the proportion of parental engagement treated children in treated and control schools (full sample); second, we regress the outcomes of the children that did not receive parental engagement on proportion of parental engagement treated children in treated and control schools (full sample). Raw scores. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A9: Spillovers (only control children - no parental engagement): Literacy skills

	Index	**Letter initial sound**	**Simple pseudo words**	**Passage reading & comprehension**
Panel A: Spillovers within schools (full sample)				
% treated parental engagement (within school)	0.864	0.036	0.861	0.974
	(0.781)	(0.744)	(0.658)	(0.818)
Observations	1740	1740	1740	1740
R^2	0.120	0.049	0.166	0.254
Unstandardized mean control	0.428	0.510	0.355	0.328
Panel B: Spillovers within control schools				
% treated parental engagement (within school)	-0.676	-1.719*	-0.026	-0.258
	(0.984)	(0.905)	(0.852)	(0.899)
Observations	834	834	834	834
R^2	0.154	0.107	0.191	0.262
Unstandardized mean control	0.428	0.510	0.355	0.328
Panel C: Spillovers within treated schools				
% treated parental engagement (within school)	0.339	-0.354	-0.189	0.970
	(0.945)	(0.673)	(0.861)	(0.854)
Observations	906	906	906	906
R^2	0.196	0.090	0.218	0.341
Unstandardized mean control	0.486	0.515	0.368	0.343

Note. The table presents the estimates for possible spillovers in treated and control schools. We proceed in two steps: first, we compute the proportion of parental engagement treated children in treated and control schools (full sample); second, we regress the outcomes of the children that did not receive parental engagement on proportion of parental engagement treated children in treated and control schools (full sample). Raw scores. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: Spillovers (only treated children - yes parental engagement): Numeracy skills

	Index	**Number write, read, compare**	**Add & subtract**	**Sequence, equations, problems**
Panel A: Spillovers within schools (full sample)				
% treated parental engagement (within school)	0.719	1.050*	0.106	0.788
	(0.712)	(0.583)	(0.719)	(0.585)
Observations	1731	1731	1731	1731
R^2	0.327	0.401	0.225	0.184
Unstandardized mean control	0.553	0.859	0.352	0.329
Panel B: Spillovers within control schools				
% treated parental engagement (within school)	0.263	1.579***	-1.034	0.309
	(0.975)	(0.467)	(1.032)	(1.203)
Observations	854	854	854	854
R^2	0.352	0.435	0.239	0.208
Unstandardized mean control	0.553	0.859	0.352	0.329
Panel C: Spillovers within treated schools				
% treated parental engagement (within school)	-1.091	-1.424**	-0.855	0.059
	(0.766)	(0.581)	(0.765)	(0.786)
Observations	877	877	877	877
R^2	0.351	0.424	0.265	0.195
Unstandardized mean control	0.514	0.832	0.332	0.301

Note. The table presents the estimates for possible spillovers in treated and control schools. We proceed in two steps: first, we compute the proportion of parental engagement treated children in treated and control schools (full sample); second, we regress the outcomes of the children that did not receive parental engagement on proportion of parental engagement treated children in treated and control schools (full sample). Raw scores. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A11: Spillovers (only treated children - yes parental engagement): Literacy skills

	Index	**Letter initial sound**	**Simple pseudo words**	**Passage reading & comprehension**
Panel A: Spillovers within schools (full sample)				
% treated parental engagement (within school)	0.481	1.052	0.215	0.120
	(0.819)	(0.693)	(0.756)	(0.692)
Observations	1731	1731	1731	1731
R^2	0.169	0.078	0.192	0.299
Unstandardized mean control	0.428	0.510	0.355	0.328
Panel B: Spillovers within control schools				
% treated parental engagement (within school)	-1.060	-0.104	-0.955	-1.279
	(0.942)	(0.806)	(0.725)	(1.051)
Observations	854	854	854	854
R^2	0.222	0.165	0.257	0.322
Unstandardized mean control	0.463	0.530	0.379	0.345
Panel C: Spillovers within treated schools				
% treated parental engagement (within school)	-2.679***	-1.011	-3.066***	-2.210***
	(0.674)	(0.744)	(0.538)	(0.732)
Observations	877	877	877	877
R^2	0.224	0.092	0.219	0.349
Unstandardized mean control	0.486	0.515	0.368	0.343

Note. The table presents the estimates for possible spillovers in treated and control schools. We proceed in two steps: first, we compute the proportion of parental engagement treated children in treated and control schools (full sample); second, we regress the outcomes of the children that did not receive parental engagement on proportion of parental engagement treated children in treated and control schools (full sample). Raw scores. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Teachers' practices on share treated parental engagement

	Index	Use teacher material	Class preparation delivery	Curriculum knowledge	Manipulative activities	Distracted students	Learning expectation
% treated parental engagement (within school)	0.090	-0.471	-0.409	0.368	0.459	0.553	-0.680
	(1.794)	(0.498)	(0.960)	(1.511)	(0.901)	(0.882)	(1.732)
Unstandardized mean control	-0.000	0.475	-0.000	-0.000	0.000	-0.000	0.000
Observations	115	115	115	115	115	115	115

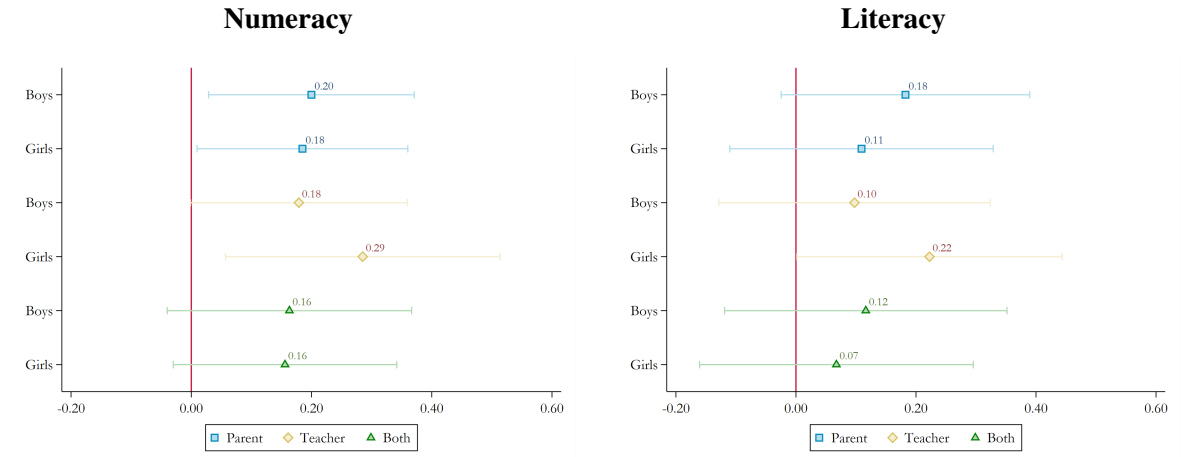
Note. Column 1 is a summary index of teacher-child interactions constructed by combining the dummy variables for the teacher practices in columns 2-7. Column 2 is the use of the teacher development material. Column 3 is class preparation/delivery (index constructed from enumerator's observations on how class is delivered). Column 4 is curriculum knowledge (index constructed from responses on whether teachers know what students should learn). Column 5 is knowledge of manipulative activities (index constructed from responses on whether teachers use manipulative activities). Column 6 is distracted students (index constructed from responses on how often students are distracted). Column 7 is learning expectations regarding students and math (index constructed from responses on what students are expected to learn). Standard errors clustered at the school level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A13: Teachers' practices on share treated parental engagement (treated schools)

	Index	Use teacher material	Class preparation delivery	Curriculum knowledge	Manipulative activities	Distracted students	Learning expectation
% treated parental engagement (within school)	4.693**	0.003	1.558*	5.104**	0.573	-0.565	2.086
	(2.077)	(0.281)	(0.766)	(2.429)	(1.553)	(1.593)	(1.339)
Unstandardized mean control	-0.000	0.475	-0.000	-0.000	0.000	-0.000	0.000
Observations	60	60	60	60	60	60	60

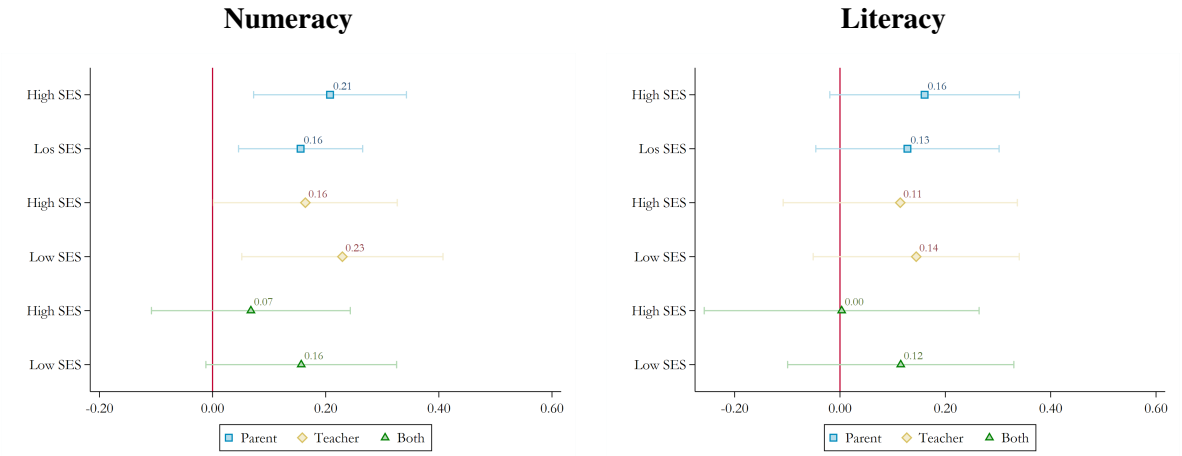
Note. Column 1 is a summary index of teacher-child interactions constructed by combining the dummy variables for the teacher practices in columns 2-7. Column 2 is the use of the teacher development material. Column 3 is class preparation/delivery (index constructed from enumerator's observations on how class is delivered). Column 4 is curriculum knowledge (index constructed from responses on whether teachers know what students should learn). Column 5 is knowledge of manipulative activities (index constructed from responses on whether teachers use manipulative activities). Column 6 is distracted students (index constructed from responses on how often students are distracted). Column 7 is learning expectations regarding students and math (index constructed from responses on what students are expected to learn). Standard errors clustered at the school level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure A1: Intention-to-treat: heterogeneity by gender



Note. The Figure presents the main estimates for numeracy and literacy skills for each treatment arm by gender. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level. Confidence interval at 95% level.

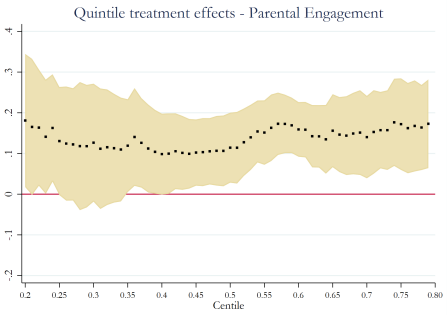
Figure A2: Intention-to-treat: heterogeneity by socio-economic status (SES)

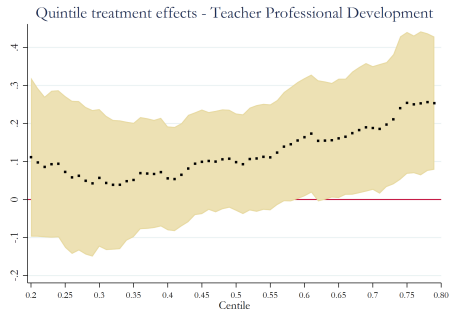


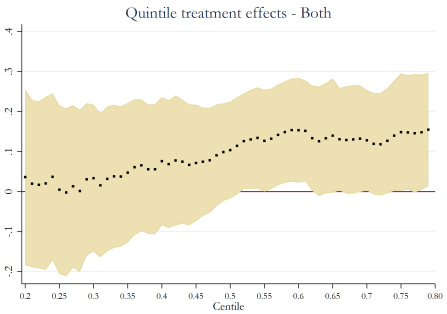
Note. The Figure presents the main estimates for numeracy and literacy skills for each treatment arm by SES status. Controls: child's gender, child's age, child's disability, parental socio-economic status, enrolment batch, wave, survey week, strata fixed effects. Standard errors clustered at the school level. Confidence interval at 95% level.

Figure A3: Quintile treatment effects

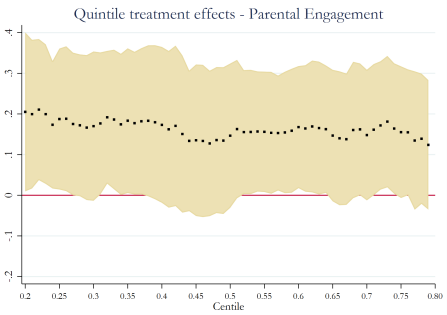
EGMA

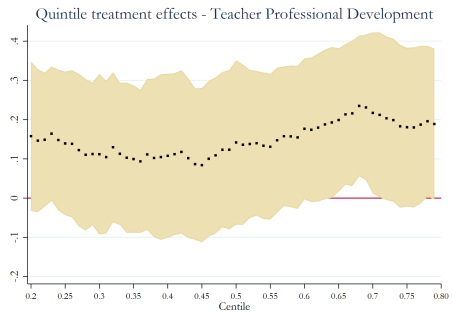


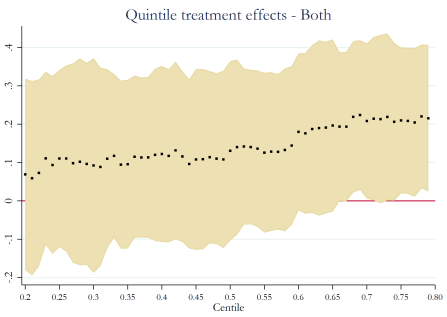




EGRA







Note. Confidence interval at 95% level.