Effects, Timing and Heterogeneity of the Provision of Information in Education: An Experimental Evaluation in Colombia

Felipe Barrera-Osorio[1]
Kathryn Gonzalez
Francisco Lagos
David Deming

This version: August, 2018

Abstract. We evaluate the effects of providing information to families on their children's reading and math achievement in a mid-size city, Colombia. Most families are poorly informed about their children's performance; our information intervention closes the gap between beliefs and performance and induces behavioral response among treated parents. We find mixed impacts of providing information on student achievement, with null findings in the first two semesters after treatment, followed by a statistically significant and positive impact and then fadeout by year two. This overall pattern is driven by large gains (and then fadeout) for students with low baseline test scores.

JEL classification codes: I21; I24; I28; O10

Keywords: education; randomization; information; parents beliefs; quality of education; Colombia

---

[1] Affiliation of all authors: Harvard Graduate School of Education. Barrera-Osorio: felipe_barrera-osorio@gse.harvard.edu; Gonzalez: kgonzalez01@g.harvard.edu; Lagos: franciscolagos@g.harvard.edu; Deming: david_deming@harvard.edu. This project was done in partnership with the LUKER Foundation, Manizales, Colombia. The commitment and professionalism of Santiago Isaza, Maria Camila Arango and Angela Trujillo, from LUKER, made this project possible. We are grateful as well with the support of the Secretary of Education of the city, and all participant public schools. We received financial support from the Weiss Family Program Fund for Research in Development Economics; from the Dean's Venture Fund Award at HGSE and from the LUKER Foundation. The study was approved by Harvard's IRB (IRB14-2842). The trial was registered at the American Economic Association archive, AEARCTR-0002192.

1. Introduction

In the human capital model, individuals invest in education if the present value of benefits exceeds costs (Becker, 1962). As such, information about the quality of education and student performance is important, since the benefits of education depend on the actual acquisition of skills in the classroom. Nonetheless, evidence suggests that families and students have limited information or are misinformed about school quality, their own academic performance, and the returns of education (Nguyen 2008, Jensen 2010, Loyalka et al. 2013). This lack of information may lead to suboptimal educational investment by households (Houtenville and Conway 2007, Avvisate et al 2014, Bergman 2015, Berlinski et al 2016, Dizon-Ross 2017). Providing performance information to parents may cause them to update their beliefs, which could lead to changes in parents' investment of time and resources in their children's education, and ultimately to increases in student achievement.

In this paper, we study the impact of providing families with standardized information about their child's own performance in school in a mid-size city in Colombia. In association with a local foundation (The Luker Foundation), we collected baseline data on the Early Grade Reading and Early Grade Math Assessments (EGRA/EGMA) of children in grades four through six in 31 public schools in the city. We visited the households of all students in the sample to collect household socioeconomic information as well as information on the beliefs of the parents about the performance of their children on EGRA/EGMA. We randomly assigned some treatment families to receive standardized information about the actual performance of their children at the end of the household interview, and we presented these families with a menu of options that they might consider in light of the information. These options ranged from asking their children about school every day to reading with their children more often. We also encouraged treatment families to strengthen their relationship with the school by interacting more with teachers.

Our results show an initial pattern of small *negative* effects (not statistically significant), then positive and significant short run effects (0.08 SD), and then zero effects in the final follow-up waves. This pattern is consistent with effects that follow a pattern of action and backsliding observed in other interventions (for instance, Gallagher, 2013 in the case of insurance, and Allcott & Rogers, 2014 in the case of electricity bills). We do not find any effect on parent internal investment. One explanation for the lack of effects on these mechanism is that all measures that proxy these variables were quite high at baseline, according to self-reported answers. We find, however, increases in the number of parents-teacher meetings and an update in the beliefs of the parents. We present some evidence suggesting that this last effect is the main channel of impact.

The second important finding of the paper is that the results are larger for students with low baseline scores (of the order of 0.27 SD, at the peak of effects). This is consistent with these students and families have less accurate information about performance, or perhaps an increase in parent-student information frictions in these households (Bergman, 2016). Still, the same pattern in the dynamics of effects is detected for this population, with a backsliding to the baseline score. We also randomly provided some teachers in the 31 schools with information about their students' academic performance. Like the intervention with parents, we encouraged teachers to engage with the families to talk about these results. However, we find no impact on student achievement of providing teachers with performance information.

In the next section, we present related literature; in Section 3 we present the description of the experiment; Section 4 discusses the data and sample. Section 5 presents the analytical plan. Section 6 shows the main results and Section 7 closes with conclusions.

2. Related Literature

Parental investment in education, namely the resources and time that is devoted by parents to support, monitor, or induce more effort in their children, has been identified as one of the main determinants of students' educational outcomes (Avvisati, Gurgand, Guyon & Maurin, 2013; Houtenville & Conway, 2008; Todd & Wolpin, 2007). Investment decisions in education critically depend on the information that is available to parents (Jensen, 2010), such that information failures may result in suboptimal investments, especially among low-income families (Dizon-Ross, 2013). Recent interventions in developed and developing countries that focus on providing information to parents have not only demonstrated positive effects on enrollment decisions and several student outcomes, including attainment and achievement in standardized tests, but also to be cost-effective (Ganimian & Murnane, 2016; Kremer, Brannen, & Glennerster, 2013). These information interventions can be broadly divided in four types, depending on what type of information is provided to parents: information about the returns to schooling (Jensen, 2010; Nguyen, 2008); information about the quality of educational institutions (Andrabi, Das & Khwaja, 2017; Banerjee, Banerji, Duflo, Glennerster & Khemani, 2010; Hastings & Weinstein, 2008); information about parenting strategies (Mayer, Kalil, Oreopoulos & Gallegos, 2015; York, Loeb & Doss, 2018); and, information about the ability, behavior, and/or academic progress of the students (Bergman, 2015; Berlinski, Busso, Dinkelman & Martinez, 2016; Dizon-Ross, 2013; Rogers & Feller, 2017). We focus here primarily on this fourth type of intervention.

In the process of making educational investment decisions, parents face at least two sources of information asymmetry. First, school staff and students themselves have information that would allow parents to make adequate educational investment decisions. But this information is not completely or timely disclosed to them due to misaligned incentives or strategic behavior from schools (Berlinski et al., 2016) or their children (Bergman, 2015). Schools may not have incentives to reveal, for example, how well a student is performing in comparison to their peers in the same school or in the city, or how to better support the student's academic progress. Second, students may be inclined to strategically disclose (hide) positive (negative) information to their parents. Interventions that provide information to parents about students' ability, behavior, or academic progress not only aim to close this information gap (Berlinski et al., 2016), but also to correct parental misbeliefs about students abilities or behavior (Dizon-Ross, 2013; Rogers & Feller, 2017) and to reduce information frictions between parents and their children's academic progress (Bergman, 2015), all of which affect educational investments.

Theoretically, once information is available to parents, they may update both their amount of effort and how it is allocated. Recent evidence confirms this hypothesis: providing information on students' ability, behavior, or academic progress not only reduces absenteeism (Berlinski et al., 2016; Rogers & Feller, 2017) and the prevalence of disruptive behaviors

(Berlinski et al., 2016), but also improves educational achievement, as measured by test scores and graduation rates (Bergman, 2015; Berlinski et al., 2016; Dizon-Ross, 2013).

Finally, while recent studies in this area have asked about longer-run effects of information provision, evidence is still scarce. Consistent with Allcott and Rogers (2014) findings on an energy conservation initiative, our results suggest an *'action and backsliding'* pattern: parents respond to information by increasing their effort, but this initial response decays unless new reports are available to them.

3. Description of the Experiment

The city of Manizales, capital of the Department of Caldas in central Colombia, is a mid-size Colombian city, with population close to half a million, and approximately 55,000 students in public basic education in grades 1 to 11. The local authorities deem education as a priority; the city has a very active civic society that is also very engaged in education policy. Authorities are interested in tackling the perceived low quality of education, as shown in national and international standardized tests. The public school system in Manizales includes 57 schools. In the present study we focus on 31 schools serving grades four, five, and six.

In association with the Luker Foundation, our study combined efforts to provide information to parents with a family engagement intervention, using an experimental design. The experimental design was divided in three phases. In the first phase, we provided report cards on school and student performance to parents of students in grades four and five. In the second phase, along with a new round of information to parents and the inclusion of more students into the experiment, we included a list of suggestions for parents on how to support their children. In the third phase, we collected follow-up information on student performance. We did not provide additional information to families during this phase[2].

Phase 1: Pilot Study (2014)

In April 2014, the Secretary of Education of Manizales (SEM) and the Luker Foundation (LF), using the Early Grade Reading Assessment (EGRA) and Early Grade Math Assessment (EGMA), collected language and math test scores of students in grades three, four and five. Relying on this information, we randomly assigned students in grades four and five into three groups: two treatment groups and one control group. We then visited households in October 2014.

In the first treatment condition (Individual student performance; Treatment Group 1), we provided families with a one-page report card that showed their child's performance, as well as their relative position to the average performance of students in the same grade and school. The information was essentially a percentile rank (e.g. 50th percentile), presented in a way that was highly salient to all families.

In the second treatment condition (School average; Treatment Group 2), we provided families with a one-page report card that showed the average score of their child's school in comparison to the average score across all schools in Manizales. Like Treatment Group 1 above, the information was a percentile rank that was translated into a reader-friendly format. In the

[2] During the third phase of the intervention we also incorporated an additional family-engagement component, focused on providing information to the teachers of students already in the experiment. This intervention led to null results. We do not report the results of that experiment here, but details of the intervention and results are provided in the Appendix.

control group, no information was provided. An example of the individual information report card provided to families in Phase 1 is presented in Figure 1.

To implement this intervention, we visited parents and guardians at their homes. These visits, which were previously scheduled by phone, were divided into three sections. First, the agent explained the objective of the study and provided the consent materials. Once the parent read the consent form, asked questions about the study and her/his participation, and signed the consent form, the visit continued. Second, the agent administered a questionnaire to the parent or guardian. One important piece of this questionnaire was to ask parents to state their beliefs for both student and school performance, using the same report card that treatment groups received, but without any information. We asked them to point to a place on a number line that represents where they thought their child and school were in the distribution of achievement. After the administration of the questionnaire concluded, the agent gave the appropriate report card to each treatment group and explained its meaning. For the control group, the visit concluded with the administration of the questionnaire. In December 2014, at the end of the academic year, we administered a new round of EGRA and EGMA tests to all students in our sample.

<u>Phase 2: Expanded Study (2015)</u>

In June 2015, SEM and LF collected again language and math test scores of all the students that were in grades three, four, or five in Phase 1. Students in grade three in Phase 1 (grade four in Phase 2) were randomly assigned into the two treatment and control groups. We modified the information treatments from Phase 1 to Phase 2 in the following ways.

In the first treatment condition (Individual student performance in the school; Treatment Group 1), the performance information for this group was the same that they received in Phase 1. However, we included in this report card a list of suggestions for parents to engage with their children's education. These included the following: recommendations on how parents can discuss their children's progress in school, and recommendations on how parents can incorporate math and literacy activities into the everyday routine.

In the second treatment condition (Individual student performance in the city; Treatment Group 2), the performance information changed from Phase 1. Instead of receiving information at the school level, this group received a one-page report card that showed their child's performance, as well as their relative position to the average performance of students in the same grade in the entire city. Also, the same list of suggestions provided to parents in Treatment Group 1 was provided to parents in Treatment Group 2. As in Phase 1, no information was provided to the control group. An example of the individual information report card provided to families in Phase 2 is presented in Figure 2; the list of recommendations provided to parents is presented in Figure 3.

We conducted a second home visit in October 2015. For students in grades four and five during Phase 1, this was their second visit. For students in grade three during Phase 1, this was their first visit. During these home visits, families in the treatment group received information regarding their children's performance on the June 2015 EGRA / EGMA tests. These home visits followed the same format we described in Phase 1. In December 2015, at the end of the academic year, we administered a new round of EGRA and EGMA tests to all students in our sample.

Phase 3: Follow-up Data Collection (2016 - 2017)

The third phase of the study included follow-up data collection for the household information intervention. In this phase, we did not provide any additional information to households or conduct additional home visits. In June 2016 and December 2016, we administered new rounds of EGRA and EGMA tests to all students in our sample who were in grades three or four in 2014. In June 2017, we administered new rounds of EGRA and EGMA tests to all students in our sample who were in grade three in 2014. We describe the timing of all data collection and study procedures in Table 1 (see Figure 4 and Appendix Table A1 for a more detailed description of the timing of the intervention and data collection).

4. Data and Sample

The analysis sample includes two cohorts of students: 3,026 students who entered the study in grades four and five in April 2014; and 1,345 students who entered the study in grade four in June 2015. On average, children in the sample were 9.9 years old and in fourth grade at baseline (April 2014 and June 2015 for the first and second cohorts, respectively), and forty-six percent of students were female. There are no significant differences between the treatment and control group on demographic characteristics (including age, gender, and grade) as well as baseline test scores collected prior to intervention (see Appendix Table A2 for details).

The primary outcomes of interest are tests from the Early Grade Reading Assessment (EGRA) and Early Grade Math Assessment (EGMA). Both the EGRA and EGMA were administered at baseline and in all follow-up waves. Student reading performance was measured using the EGRA, and student scores were based on the number of words correctly read. Student math performance was measured using the EGMA, and student scores were based on the number of subtraction problems solved correctly.[3] For our main analyses, we create a composite measures of student achievement by combining standardized scores across the math and reading assessments (see Table A3 for the correlations between student performance over time and between subjects).

The percent of students with outcome information differed across waves, ranging between 69 percent and 89 percent. However, we do not see any evidence of differential test-taking based on treatment status in any of the follow-up waves. We also do not find evidence that baseline test score is related to the interaction between missing follow-up test score and treatment status, across all follow-up waves (see Appendix Table A4 for details).

Parental beliefs: At each home visit, we elicited parents' beliefs about their children's performance on the EGRA and EGMA assessments. We asked parents about the total number of words they expected their children to read correctly on the EGRA assessment, and the total number of subtraction problems they expected their children to solve correctly on the EGMA assessment. We also asked parents their beliefs about the average number of words read and subtractions problems solved correctly among other children attending their child's school.

[3] The EGMA assessment also includes two additional components: sums and problems. However, these outcomes were not collected for all grades across all outcome waves. We also observed moderate ceiling effects for both of these measures (see Appendix Figures A1 and A2). Therefore, we focus on subtractions as our measure of mathematics).

Parental behaviors: At each home visit, we also asked parents about their behaviors around investing in the children's education. Specifically, we asked parents to state the number of days per week, on average, they engaged in the following activities with their child: asking about school, helping with studying, reading to him or her, helping with homework, and asking about grades. We also collected information about parents' relationship with the school, including how frequently parents attended guardians' meetings, parents' school, school activities, and meetings with teachers (Never/Almost never/Occasionally/Almost always/Always).

We collected information on the characteristics of participating families during the initial home visits. On average, the responding parent or guardian was approximately 39 years old. In nearly ninety percent of households the father was reported as working, while the mother reported working in just under 50 percent of households. Both mothers and fathers had approximately eight years of education, and average household income was approximately two minimum salaries (see Appendix Table A5 for details).

Parents also reported relatively high involvement in their children's education at the initial home visit prior to the intervention. On average, parents reported asking about school, helping with homework, and asking about grades nearly five days per week. Parents reported helping their children study approximately four days per week, and reported reading to their children approximately three days per week. There are few significant differences in household characteristics or parent investment behavior prior to intervention between the treatment and control group. Parents of children in the treatment group reported asking about school and helping their children with studying more frequently relative to parents in the control group (differences of 0.07 days and 0.15 days, respectively) but were not more likely to report reading with their children, helping with homework, or asking with grades (see Appendix Table A5 for details).

We also examine the extent to which parents' beliefs regarding their children's performance on the math and reading assessments accurately reflected their children's actual performance. We asked parents to state the score they expected their children to receive on initial EGRA and EGMA assessments, and find large differences between students' raw scores and their parents' beliefs. These results are reported in Table 2. On average, parents overestimated their children's reading performance by 10.5 points (a difference of nearly 0.5 SD). However, we see a very different pattern for math. On average, underestimated their children's performance on the subtractions assessment by 5.2 points (a difference of nearly 1 SD).

We also find some evidence that parents' accuracy of beliefs regarding baseline performance is related to parental education. Parents with higher levels of education had more accurate beliefs regarding students' reading performance at baseline after controlling for other household characteristics, although there were no differences in the accuracy of parents' beliefs based on education for math performance (see Table 2).

5. Analytic Models

For our main analysis, we estimate a series of intent-to-treat (ITT) empirical models which provide a set of causal estimates of the effect of providing parents with information on their children's academic performance on children's subsequent achievement. The ITT estimates measure the effect of being assigned to the treatment condition, in which parents were assigned to receive a home visit during which they received a report containing information on school and student performance. Specifically, we estimate models of the following form:

$$(1) \qquad Y_{ij} = \alpha + \beta Treatment_i + \gamma X_i + \epsilon_{ij}$$

The variable Y_{ij} represents the test score for student *i* in school *j*. This is regressed on the variable $Treatment_i$ which is an indicator for whether the household of student *i* was assigned to either Treatment Group 1 or Treatment Group 2. We also include a series of child baseline covariates, X_i, which include student age, gender, baseline math test scores and baseline reading test scores, and grade in 2014. We estimate a model for test scores collected in each of the follow-up waves between December 2014 and June 2017. In all models, we cluster standard errors at the school level. For analyses of impacts on EGRA/EMGA scores in the follow-up waves, we cluster standard errors based on the school attended by the student in each follow-up wave. For analyses of impacts on parent beliefs and behaviors, we cluster standard errors based on the school the student attended at baseline since not all students who received a second home visit could be located at the school during the administration of the EGRA/EGMA in June or December 2015.

To examine the extent to which there are heterogeneous impacts based on students' baseline achievement, we estimate models of the following form that include an interaction between treatment status and baseline achievement:

$$(2) \quad Y_{ij} = \alpha + \beta Treatment_i + \delta Treatment_i * LowAchievement_i + \\ + \rho LowAchievement_i + \gamma X_i + \epsilon_{ij}$$

The variable $LowAchievement_i$ is an indicator for whether student i scored low on either the baseline math or reading assessments (i.e., below the 25^{th} percentile)[4]. We estimate analogous models to examine whether impacts differ by parents' initial beliefs of student ability were above or below parents' beliefs regarding average school performance.

We use dummy variable adjustment to account for missing baseline covariates. In cases where students were missing baseline math and/or reading scores, we set the missing values to the overall mean. For each subject, we created indicator variables set to one if the baseline score was missing and zero otherwise. These indicator variables were included in all analyses.

6. Results

a. Impacts of information receipt on parental beliefs and behaviors

We first examine whether parents' beliefs about their children's math and reading performance appeared to be affected by the receipt of information one year after receiving the initial report card. If the parents' beliefs were not affected by receipt of the report card information, it is unlikely we would observe changes in parent behavior and subsequent impacts on student

[4] For the distributions of baseline scores and the 25^{th} percentile cutoff for reading and subtractions, see Figures A5 and A6.

performance. In the second home visit, as in the first home visit, we elicited parents' beliefs about their children's performance on the math and reading assessments.

In Table 3, we confirm that there is evidence that the beliefs of parents in the second home visit among treated household differed from those of parents in control households. We see that on average, parent beliefs regarding their children's math performance in the treated group were lower relative to parent beliefs in the control group. As parents' initial beliefs regarding their children's baseline math performance were higher relative to the children's actual baseline performance, this suggests that parents updated their beliefs in response to the information provided in the intervention. Although we see no overall difference with respect to parents' beliefs about the number of words their children read correctly, we find that the extent to which the treatment shifted parents' reading beliefs differed based on students' baseline performance. While parents shifted their beliefs regarding reading upwards for students who scored at or above the 25^{th} percentile in reading, parents shifted their beliefs downward if their children were below the 25^{th} percentile. We also find that the gap between parent beliefs in the second home visit and their children's actual performance in the previous follow-up wave (June 2015) were smaller in the treatment group relative to the control group.

Table 4 presents effects of the program on parents' educational investment behavior. Parents in the treatment group were 7.3 percentage points more likely to report consistently (always) attending meeting with teachers relative to the control group. However, they do not observe effects on other aspects of parents' relationship with the school. The treatment also did not have a significant impact on a range of other parent behaviors regarding families' internal investment in education, including the number of days per week parents reported engaging in the following activities: asking about school, helping with studying, reading with their child, helping with homework, and asking about grades (see Appendix Table A5 for details).

b. Impacts of information receipt on student performance

Table 5 presents effects of information on math and reading scores across all follow-up waves. There was no statistically significant impact in either December 2014 and June 2015, which represent approximately two months and eight months after treatment, respectively, for the first treated cohort. However, we find evidence of small, positive impacts in the December 2015 (0.08 SD) and the June 2016 (0.11 SD) follow-up waves, although the latter impact is not statistically significant. These represent impacts approximately 14 months and 20 months after treatment for the first cohort, and two months and 8 months after treatment for the second cohort. However, these impacts do not persist through the final two follow-up waves of the study.

It is important to note that the estimates presented for each of the follow-up waves presented in Table 5 represent the average impact across three cohorts of students which differed in the timing and duration of treatment. First, different grade cohorts exited and entered the study at different points. For example, students who were in grade five in 2014 were not followed after December 2015; students who were in grade three in 2014 did not enter the study until June 2015. Furthermore, as described above, home visits occurred during October 2014 and October 2015. As a result, students in the treatment group who were in grades four and five in 2014 received the first home visit in October 2014 and a second home visit in October 2015. The impact in the December 2015 follow-up wave therefore represents the combined effect of receiving information during two separate home visits over the course of approximately one year. In contrast, the first home visit for students who were in grade three in 2014 occurred in October 2015; therefore, the impact in the December 2015 follow-up wave represented the effect of the intervention roughly two months after the start of the intervention. If the effects of receiving information on students' academic outcomes increases over time (for example, if parents are not able to immediately respond to the information provided by the information, but are able to do so over time), we would expect to see differences in the trajectories of impacts across the three grade cohorts. Therefore, the overall impact in each follow-up wave could mask heterogeneity by grade cohort.

To examine whether this is the case, we estimate our main specification separately for students in grades three, four, and five in 2014. As shown in Table 6 and Figure 5, the patterns of results across the follow-up waves differ across the three grade cohorts. Among students who were in grade five in 2014, impacts are increasing over time. There is a small, negative and imprecisely estimated impact in the December 2014 follow-up wave. However, impacts are positive in the June 2015 follow-up wave (0.11 SD). Among this cohort, the largest impact occurs in the final follow-up wave for which we observe this cohort (December 2015; 0.15 SD), which occurred shortly after the second home visit. Among the grade four cohort, impacts are less positive. Impacts in the first two waves are negative, although not statistically significant. Point estimates in the later follow-up waves are imprecisely estimated and mixed in sign. Finally, students in the grade three cohort show evidence of positive impacts in the second follow-up wave following the home visit (0.11 SD), but no evidence of impacts in later waves. As the study did not include additional intervention in either 2016 or 2017, this suggests that the short-term impacts of providing information on students' academic performance decrease out over time in the absence of continued intervention ("backsliding").

c. Heterogeneity by baseline student performance and by baseline parental beliefs

Next, we explored whether treatment impacts differ based on students' baseline math and reading performance. As described above, there is some evidence that the extent to which the provision of information affected parents' beliefs about their children's math and reading performance differed based on baseline student performance. Therefore, it is possible that impacts on student outcomes similarly differed based on baseline performance. Our results presented in Table 7 suggest that the despite the overall null impacts across a majority of follow-up waves, the information treatment may have benefited the lowest-performing students. Impacts are positive in three of the five follow-up waves for students who scored below the 25th percentile on either the baseline math or reading assessment, with effects ranging from 0.10 SD to 0.27 SD. As shown in Figure 6, the treatment closed the gap between students in grade 4 in 2014 who were relatively high- and low-performing at baseline in the follow-up waves following the home visits; however, the gap widened in later follow-up waves[5].

Finally, we examine whether treatment impacts on students' math and reading performance differ based on the accuracy of parents' initial beliefs regarding their children's performance. We find that impacts on both the math score and composite math and reading score did not vary based on the absolute value of the gap between parents' beliefs and performance at baseline (see Appendix Tables A8 and A9 for details).

[5] We also examine the sensitivity of this result to other ways of classifying students as relatively high performing and low performing at baseline. For the full results, see Appendix Tables A12 through A14.

However, parents' beliefs regarding the raw number of correct reading or subtractions may not be the most salient measure of parents' beliefs regarding their children's academic achievement. We therefore also examine whether impacts differ based on parents' perception of their children's performance relative to the performance of their peers. In the initial home visit, parents were asked about their beliefs both about the average performance of students in their child's school on the math and reading assessments, in addition to the performance of their own child. Whether parents placed their child above or below the school average is another potential indicator of parents' perception of their child's performance. Therefore, we examine whether impacts vary based on whether the parents' beliefs regarding their child's math and/or reading score was below parents' beliefs regarding the school average. As shown in Table 8, we find some evidence that impacts are more negative where parents believed their child's baseline math and/or reading performance was below that of other children in their schools. In the June 2016 and December 2016 follow-up waves, estimated impacts were lower by 0.31 SD and 0.32 SD, respectively, in households where parents believed their child's math or reading performance was below the school mean.

d. Heterogeneity by the interaction between baseline parental beliefs and student performance

In light of the results presented above, we consider whether the effect of parents' receipt of information varies based on the interaction between parents' baseline beliefs and students' baseline performance. Specifically, we estimate a model that includes an intervention between treatment status, baseline parent beliefs, and baseline student achievement. A positive coefficient on the triple interaction indicates that parent beliefs moderate the negative relationship between treatment impact and baseline achievement described above.

The results of estimating this model are presented in Table 9. We find no evidence that the moderating effect of baseline achievement on the effect of the intervention on reading differs based on parent beliefs. However, we do find evidence that the triple interaction between status, baseline achievement, and parents' baseline beliefs has a significant association with math performance. The estimate of this effect is positive and significant across four of the six follow waves, indicating that the impact of the treatment was larger when parent baseline beliefs and students' baseline achievement were more in alignment. Specifically, for students with lower baseline achievement the treatment impact was larger when initial beliefs were higher; for students with higher baseline achievement the effect is larger when initial baseline beliefs were lower.

7. Conclusion

Parents have limited and, sometimes, erroneous information about the academic performance of their children. As such, information failures may induce misallocation of resources and suboptimal investment in education—either within the interior of the household (e.g., time, resources) or in the relationship with the school. In this intervention we aimed to solve the problem of information by providing results in an early assessment on literacy and math.

We show that parents do have erroneous information about the academic performance of their children. Upon receiving information, parents update their assessment of the performance of their children and meet more frequently with the teachers. Also, we show that the provision of information lead to some improvements in the academic performance of children, especially students with low scores at baseline. We demonstrate that families can react to information, and that the provision of information may be one leverage to increase learning. However, these effects are short-lived. After some time, the control group catches up with the group, producing a dynamic of “action” (in the short run) and backsliding to the (control) mean. This is consistent with models in which the parents cannot permanently modify their behavior (or change a “stock” or permanent variable); parents can temporarily modify their behavior in the short run but actions quickly return to a “business as usual” mode. Thus, either information must be provided very frequently, or information alone is insufficient to permanently alter performance levels.

References

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, *104*(10), 3003-37.

Andrabi, T., Das, J., & Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, *107*(6), 1535-63.

Avvisati, F., Gurgand, M., Guyon, N., & Maurin, E. (2013). Getting parents involved: A field experiment in deprived schools. *Review of Economic Studies*, *81*(1), 57-83.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R., & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, *2*(1), 1-30.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2), 9-49.

Bergman, P. (2015). Parent-child information frictions and human capital investment: Evidence from a field experiment. Mimeo, Columbia University. http://www.columbia.edu/~psb2101/BergmanSubmission.pdf

Berlinski, S., Busso, M., Dinkelman, T., & Martinez, C. (2016). Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile. *Unpublished Manuscript*.

Dizon,-Ross, R. (2017). Parents Beliefs About Their Children's Academic Ability: Implications for Educational Investments. Mimeo, University of Chicago, Booth School of Business, http://faculty.chicagobooth.edu/rebecca.dizon-ross/research/papers/perceptions.pdf

Dizon-Ross, R. (2013). *Parents' perceptions and their children's education: Experimental evidence from Malawi*. Working Paper. Export BibTex RTF Tagged XML RIS perceptions_paper_2013dec4. pdf 4.35 MB Admin Login OpenScholar.

Fryer Jr, R. G. (2016). Information, non-financial incentives, and student achievement: Evidence from a text messaging experiment. *Journal of Public Economics*, *144*, 109-121.

Gallagher, J. (2014). Learning about an infrequent event: evidence from flood insurance take-up in the United States. *American Economic Journal: Applied Economics*, *6*(3), 206-33.

Ganimian, A. J., & Murnane, R. J. (2016). Improving education in developing countries: Lessons from rigorous impact evaluations. *Review of Educational Research*, *86*(3), 719-755.

Hastings, J. S., & Weinstein, J. M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics*, *123*(4), 1373-1414.

Houtenville, A. J., & Conway, K. S. (2008). Parental effort, school resources, and student achievement. *Journal of Human resources*, *43*(2), 437-453.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515-548.

Kremer, M., Brannen, C., & Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, *340*(6130), 297-300.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., Zhang, L., Shi, Y., Chu, J., & Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? Evidence from China. *Journal of Comparative Economics*, 41, 1012–1025.

Mayer, S. E., Kalil, A., Oreopoulos, P., & Gallegos, S. (2015). *Using behavioral insights to increase parental engagement: The parents and children together (PACT) intervention* (No. w21602). National Bureau of Economic Research.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from Madagascar. *Unpublished manuscript*.

Rogers, T., & Feller, A. (2017). Reducing Student Absences at Scale By Involving Families. *Unpublished manuscript*.

Todd, P. E., & Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, *1*(1), 91-136.

York, B. N., Loeb, S., & Doss, C. (2018). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. *Journal of Human Resources*, 0517-8756R.

Tables and Figures

Table 1. Timing of intervention with percent of students with outcome information in each wave

	April 2014	Dec 2014	June 2015	Dec 2015	June 2016	Dec 2016	June 2017
Individual information							
Intervention Group A (Grade 5 in 2014)	1,606 (100 %)	1,422 (89%)	1,356 (84%)	990 (62%)	--	--	--
Intervention Group B (Grade 4 in 2014)	1,420 (100 %)	1,282 (90%)	1,237 (87%)	954 (67%)	1,213 (85%)	1,171 (82%)	--
Intervention Group C (Grade 3 in 2014)	--	--	1,345 (100%)	1,053 (78%)	1,205 (90%)	1,165 (87%)	1,036 (77%)

Table 2. Difference between parent beliefs and performance at baseline (actual performance – beliefs)

	Mean	SD	N	Gap (mean)	Abs(Gap) (mean)	N
Number of words read	91.91	20.09	4,362	10.51	19.59	2,037
Number of correct subtractions	12.75	5.29	4,190	-5.21	7.43	1,988

Relationship between accuracy of beliefs and parent education

	Abs(Gap): Reading		Abs(Gap): Subs	
Mother's years of education	-0.596***	-0.381***	0.006	-0.002
	(0.130)	(0.138)	(0.035)	(0.038)
Observations	1,889	1,861	1,843	1,816
Dependent variable mean	19.591	19.509	7.362	7.386
Includes child and parent controls	No	Yes	No	Yes

Baseline performance based on administration of EGRA/EGMA assessment in April 2014 among students who were in grades 4 or 5 in 2014, and assessment in June 2015 among students who were in grade 3 in 2014. Parent beliefs based on home visit conducted in October 2014 among students who were in grades 4 or 5 in 2014, and in October 2015 among students who were in grade 3 in 2014. The gap between parent beliefs and performance is calculated by the difference between parent beliefs regarding student performance (e.g. number of words read, correct subtractions, etc.) and student performance at baseline (e.g. number of words read, correct subtractions, etc.). Child and parent controls include child age and gender, mother's occupational status, and household income (less than one MS, one MS, between one and two MS, at least two MS). Standard error in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p<0.01$.

Table 3. Impact on parent beliefs

	Post-intervention beliefs		Gap between post-intervention beliefs and June 2015 performance (absolute value)	
Number of words read				
Any treatment	2.115	4.009**	-3.512**	-3.553**
	(1.470)	(1.738)	(1.417)	(1.593)
Treatment*Low student reading performance at baseline		-7.973***		0.092
		(2.394)		(2.983)
Observations	1962	1957	1965	1864
Number of correct subtractions				
Any treatment	-0.372*	-0.347	0.291	0.214
	(0.184)	(0.242)	(0.230)	(0.228)
Treatment*Low student math performance at baseline		0.243		-0.224
		(0.455)		(0.514)
Observations	1962	1957	1965	1864

Low baseline performance indicates that student was below the 25th percentile on math and/or reading at baseline. Baseline performance based on administration of EGRA/EGMA assessment in April 2014 among students who were in grades 4 or 5 in 2014, and assessment in June 2015 among students who were in grade 3 in 2014. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. Impact on parents' relationship with the school

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Guardians' meetings		Parents' school		School activities		Meetings with teachers	
Treatment	-0.002	-0.015	0.004	-0.005	0.015	0.008	0.073***	0.090***
	(0.010)	(0.018)	(0.016)	(0.022)	(0.027)	(0.027)	(0.019)	(0.029)
Treatment * Low baseline math and/or reading performance		0.027		0.029		0.040		-0.042
		(0.030)		(0.047)		(0.031)		(0.053)
Observations	1970	1894	1970	1894	1970	1894	1970	1894
Control mean	0.89		0.55		0.66		0.68	

All outcomes are binary indicators for whether parents reported "Always" when asked how frequently they participated in each activity. Other possible responses included "Almost always", "Occasionally," "Almost never," and "Never." All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses. Standard errors clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Impact on composite test score outcomes

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Math and reading composite						
Treatment	-0.035	0.010	0.083**	0.106	-0.078	-0.020
	(0.024)	(0.045)	(0.035)	(0.068)	(0.099)	(0.114)
Observations	2661	2593	2994	2418	2336	1036
Control mean	0.451	1.288	1.334	2.740	2.422	2.208

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group at baseline. The sum of the standardized reading and math scores were calculated for each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

First and second columns: Includes students who were enrolled in grades 4 or 5 in 2014. Treatment group includes students who were assigned to receive information on individual performance *and* students assigned to receive information on school performance. Students assigned to receive information on school performance receive information on individual performance in subsequent waves. Control group includes students who were assigned to the control condition.

Third column: Includes students who were enrolled in grades 3, 4 or 5 in 2014. Treatment group includes students who were assigned to receive information on individual performance in 2014 and/or 2015. Control group includes students who were assigned to the control condition.

Fourth and fifth columns: Includes students who were enrolled in grades 3 or 4 in 2014. Treatment group includes students who were assigned to receive information on individual performance in 2014 and/or 2015. Control group includes students who were assigned to the control condition.

Sixth columns: Includes students who were enrolled in grade 3 in 2014. Treatment group includes students who were assigned to receive information on individual performance in 2015. Control group includes students who were assigned to the control condition.

Table 6. Separately by grade – Using all students in each grade

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Grade 5 in 2014 (Cohort 1)						
Treatment	-0.022	0.114**	0.151**	--	--	--
	(0.040)	(0.051)	(0.059)			
Observations	1393	1356	989			
Grade 4 in 2014 (Cohort 1)						
Treatment	-0.050	-0.121	0.073	0.096	-0.061	--
	(0.042)	(0.075)	(0.069)	(0.131)	(0.131)	
Observations	1268	1237	952	1213	1171	
Grade 3 in 2014 (Cohort 2)						
Treatment	--	--	0.024	0.112*	-0.093	-0.020
			(0.050)	(0.064)	(0.097)	(0.114)
Observations			1053	1205	1165	1036
Overall impact, pooling across all grades						
Treatment	-0.035	0.010	0.083**	0.106	-0.078	-0.020
	(0.024)	(0.045)	(0.035)	(0.068)	(0.099)	(0.114)
Observations	2661	2593	2994	2418	2336	1036

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group at baseline. The sum of the standardized reading and math scores were calculated for each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7. Impact on composite test score outcomes, by baseline performance

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Math and reading composite						
Any treatment	-0.121***	-0.053	-0.017	0.013	-0.081	0.001
	(0.039)	(0.047)	(0.040)	(0.079)	(0.087)	(0.126)
Treatment*Low baseline performance	0.222***	0.134	0.206**	0.255	-0.011	-0.051
	(0.065)	(0.096)	(0.079)	(0.161)	(0.151)	(0.229)
Impact on low-performing students	0.101**	0.081	0.189***	0.268**	-0.092	-0.050
Observations	2510	2445	2894	2367	2290	1036

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated, and standardized again with respect to the control group, to form the composite score. Low baseline performance indicates that student was below the 25th percentile on math and/or reading at baseline. Omitted category is students who scored at or above the 25th percentile on both math or reading at baseline. Baseline performance based on administration of EGRA/EGMA assessment in April 2014 among students who were in grades 4 or 5 in 2014, and assessment in June 2015 among students who were in grade 3 in 2014. All models include controls for age, gender, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
First and second columns: Includes students who were enrolled in grades 4 or 5 in 2014. Treatment group includes students who were assigned to receive information on individual performance *and* students assigned to receive information on school performance. Students assigned to receive information on school performance receive information on individual performance in subsequent waves. Control group includes students who were assigned to the control condition.
Third column: Includes students who were enrolled in grades 3, 4 or 5 in 2014. Treatment group includes students who were assigned to receive information on individual performance in 2014 and/or 2015. Control group includes students who were assigned to the control condition.
Fourth and fifth columns: Includes students who were enrolled in grades 3 or 4 in 2014. Treatment group includes students who were assigned to receive information on individual performance in 2014 and/or 2015. Control group includes students who were assigned to the control condition.

Table 8. Impact on composite outcomes, by baseline parent beliefs

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Treatment	-0.136	0.018	0.083	0.373***	0.116	0.121
	(0.086)	(0.107)	(0.096)	(0.109)	(0.139)	(0.175)
Treatment* Parent beliefs – Student math and/or reading score below school mean	-0.057	-0.112	-0.099	-0.309*	-0.322*	-0.238
	(0.141)	(0.239)	(0.144)	(0.165)	(0.169)	(0.234)
Observations	992	966	1189	1379	1327	765

Parent belief variable is an indicator for whether parent beliefs regarding the number of words read and/or correct subtractions by the student is greater than parent beliefs regarding the average number of words read/correct subtractions in the student's school. Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated for baseline and each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9. Triple interaction between treatment, beliefs, and achievement – Separately by subject

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Number of correct words read (unstandardized)						
Treatment	-0.075	1.816	-0.471	0.832	-1.801	0.710
	(1.259)	(1.737)	(0.507)	(1.856)	(1.701)	(2.468)
Baseline achievement (# correct, mean-centered)	0.739***	0.663***	0.370***	1.025***	0.929***	0.449**
	(0.109)	(0.112)	(0.060)	(0.123)	(0.112)	(0.179)
Baseline beliefs (# correct, mean-centered)	-0.067	-0.052	0.034	0.187***	0.106	-0.231**
	(0.043)	(0.068)	(0.024)	(0.069)	(0.090)	(0.097)
Baseline achievement*Beliefs	0.000	0.003	-0.003	0.004	-0.000	-0.009
	(0.002)	(0.003)	(0.002)	(0.004)	(0.005)	(0.010)
Treat* Baseline achievement	-0.192**	-0.164	-0.124**	-0.275*	-0.135	-0.067
	(0.094)	(0.113)	(0.052)	(0.162)	(0.139)	(0.228)
Treat* Baseline beliefs	0.079*	0.126	-0.050*	-0.223***	-0.062	0.255**
	(0.043)	(0.079)	(0.026)	(0.073)	(0.098)	(0.105)
Treat*Achievement*Beliefs	-0.000	-0.005	0.003	-0.007	0.005	0.003
	(0.003)	(0.003)	(0.002)	(0.005)	(0.006)	(0.012)
Observations	985	949	1513	1364	1312	764
Number of correct subtractions (unstandardized)						
Treatment	-0.614	0.035	0.086	0.290	-0.376	-0.313
	(0.523)	(0.314)	(0.253)	(0.249)	(0.322)	(0.408)
Baseline achievement (# correct, mean-centered)	0.555***	0.345***	0.396***	0.322***	0.275***	0.288***
	(0.121)	(0.069)	(0.049)	(0.055)	(0.051)	(0.084)
Baseline beliefs (# correct, mean-centered)	0.117*	-0.017	0.052	-0.008	-0.026	-0.012
	(0.064)	(0.074)	(0.056)	(0.038)	(0.052)	(0.048)
Baseline achievement*Beliefs	-0.003	-0.012	-0.015**	-0.021***	-0.025**	-0.025**
	(0.010)	(0.012)	(0.006)	(0.007)	(0.010)	(0.011)
Treat* Baseline achievement	-0.113	-0.036	0.003	0.074*	0.084	-0.022
	(0.131)	(0.062)	(0.072)	(0.044)	(0.058)	(0.104)
Treat* Baseline beliefs	-0.028	0.034	-0.008	0.018	0.068	0.064
	(0.076)	(0.073)	(0.060)	(0.049)	(0.062)	(0.056)
Treat*Achievement*Beliefs	-0.003	0.012	0.013*	0.025**	0.026**	0.022*
	(0.013)	(0.012)	(0.008)	(0.010)	(0.011)	(0.013)
Observations	939	910	1484	1361	1309	762

All models include controls for age, gender and grade. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1. Individual information report card in Phase 1

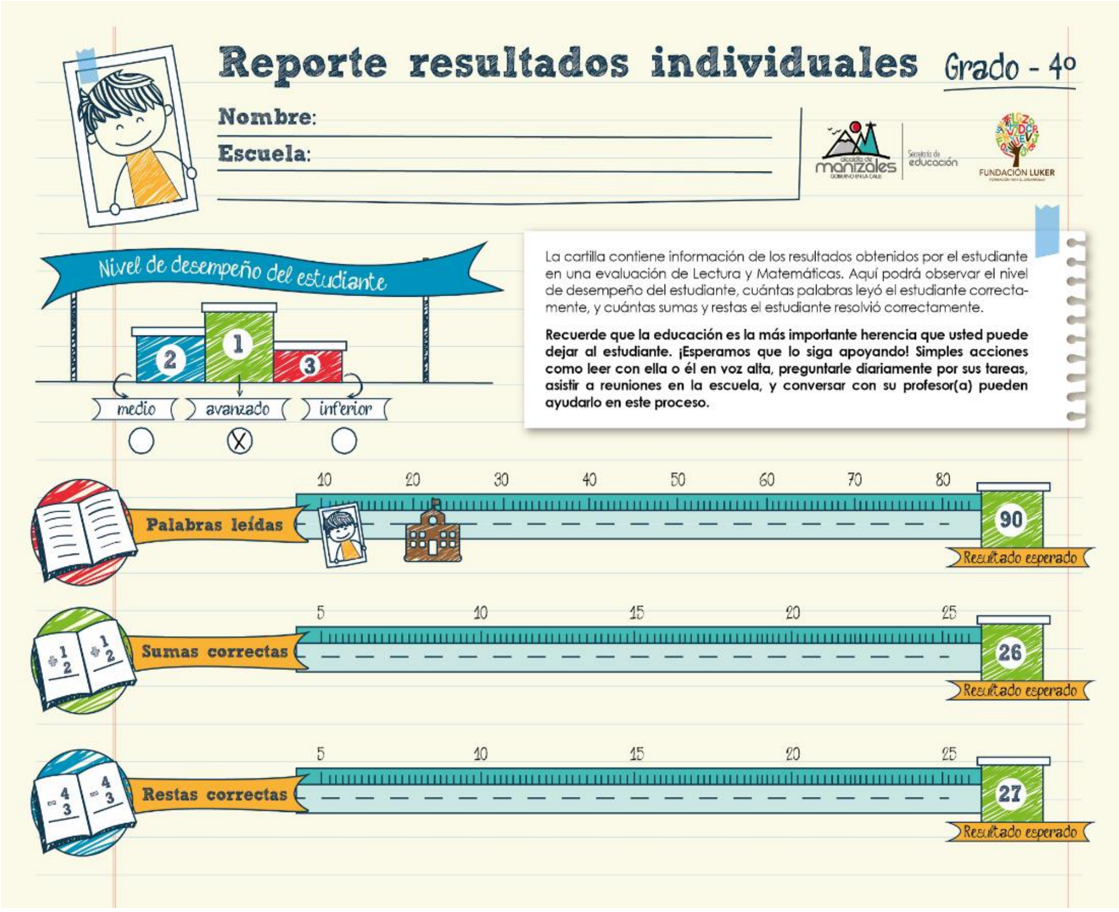

Figure 2. Individual information report card in Phase 2

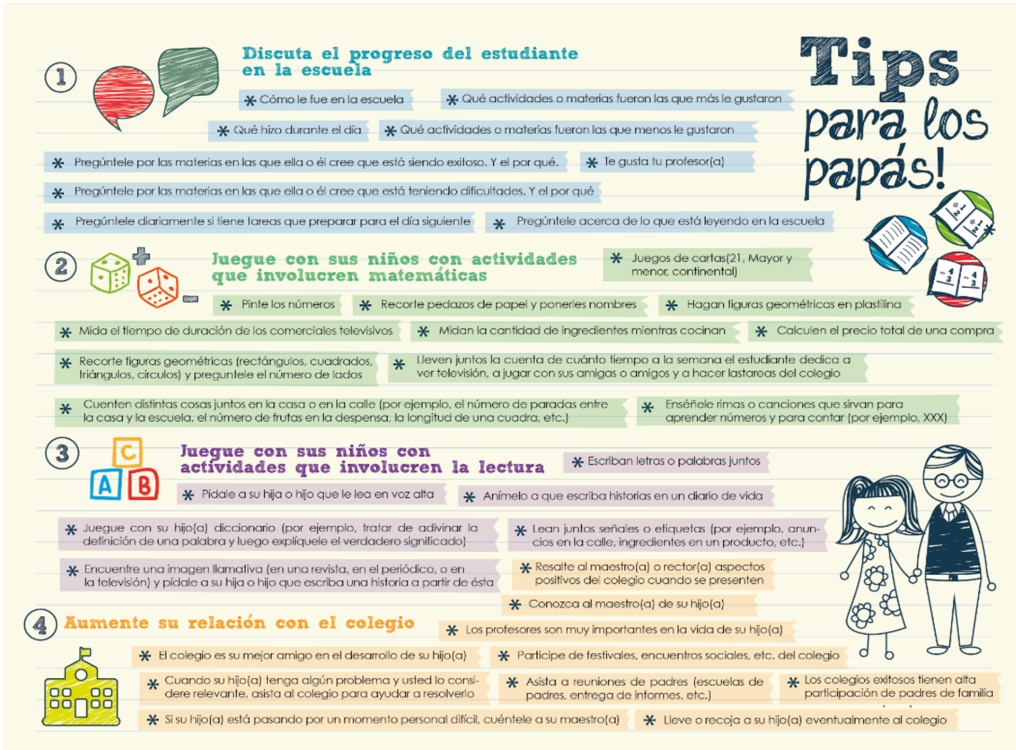

Figure 3. Recommendations for parents and teachers provided in Phase 2 and Phase 3

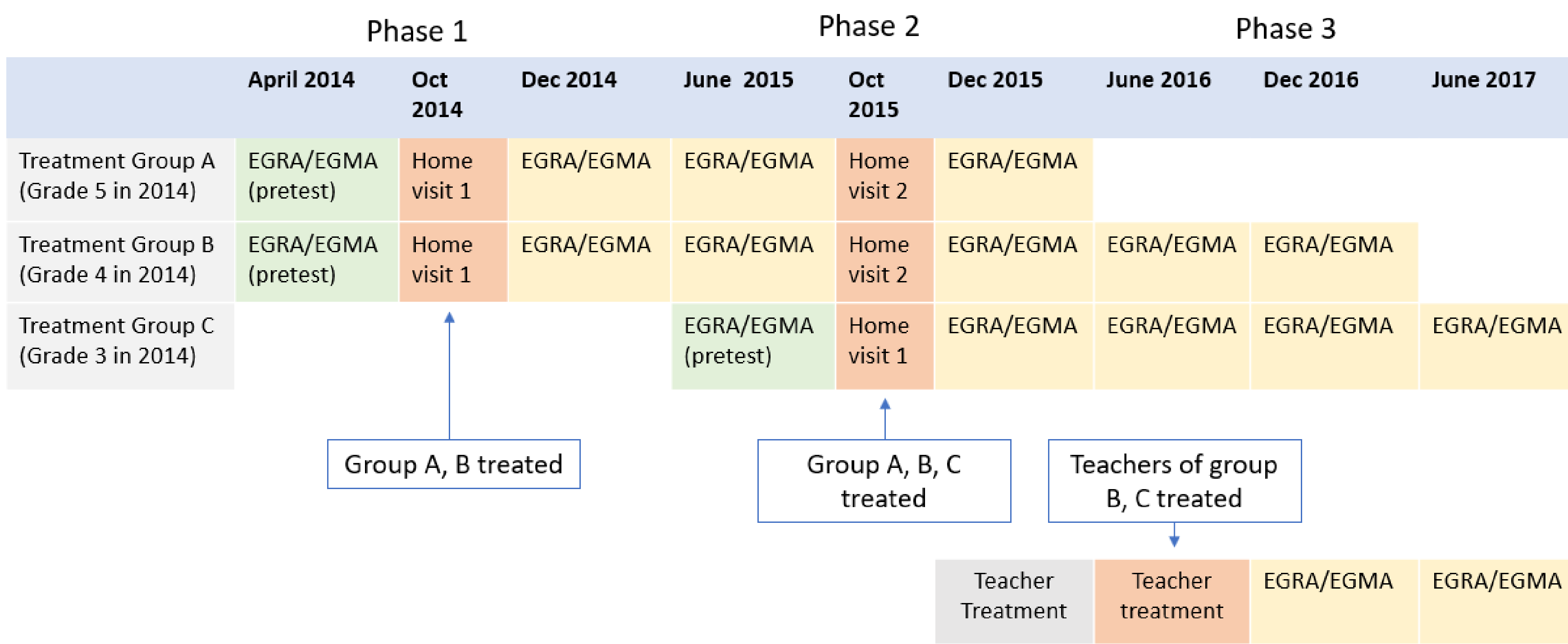

Figure 4. Overview of intervention timing

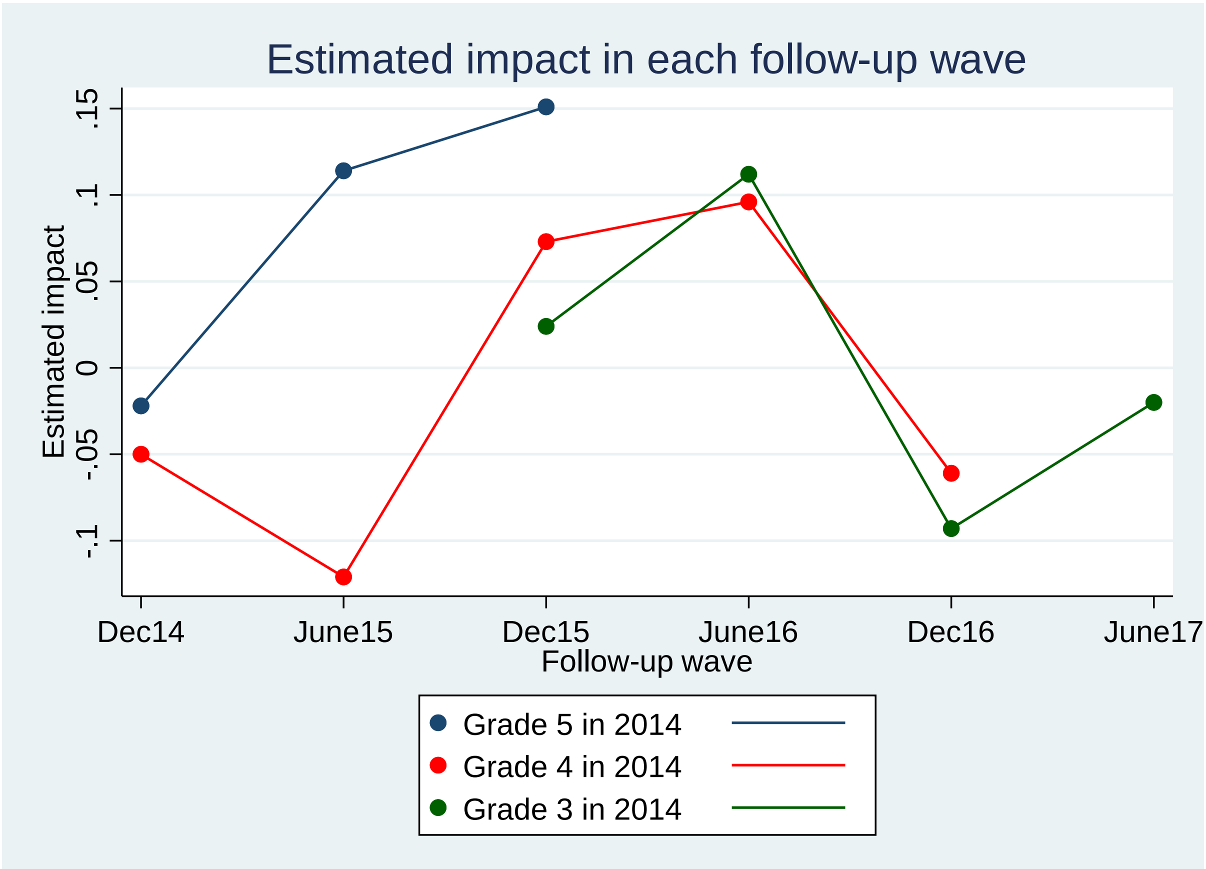

Figure 5. Estimated treatment impact across each follow-up wave, by grade**.**

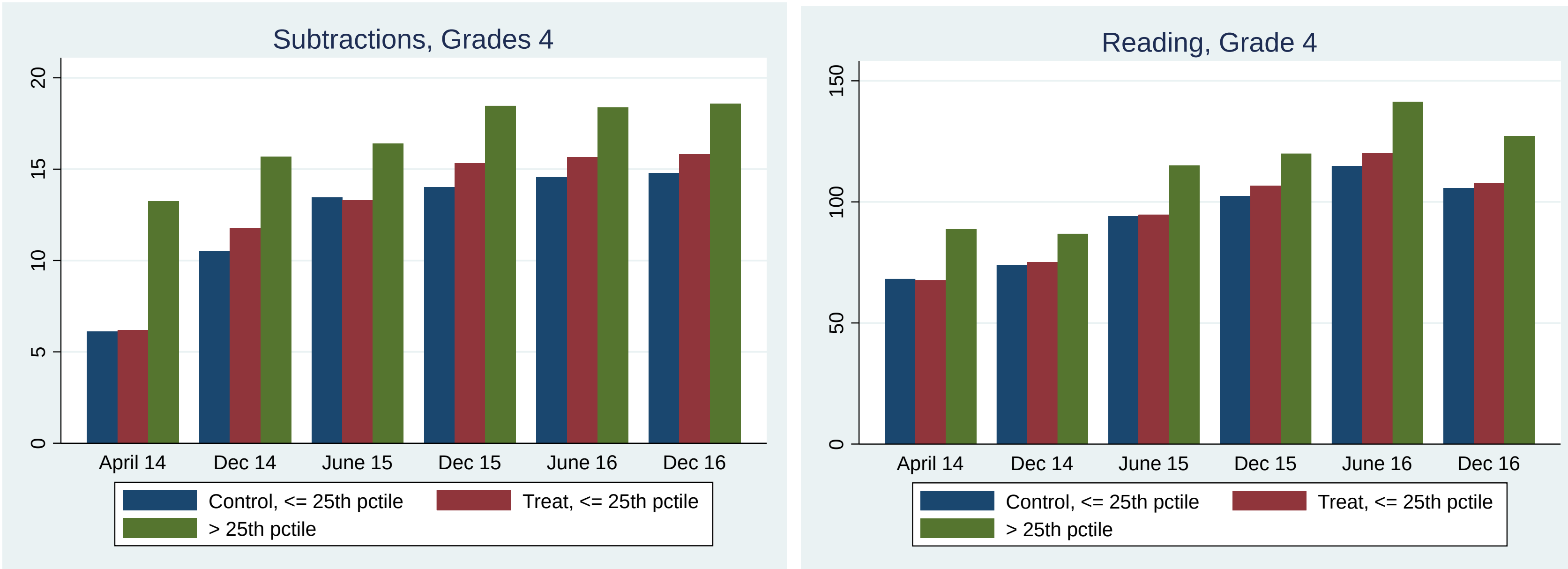

Figure 6. Reading and subtractions scores in each wave for students in grade 4 in 2014. Figures present unstandardized means for students in the following groups: Control group and below the 25th percentile at baseline; treatment group and below the 25th percentile at baseline; treatment and control group (combined) above the 25th percentile at baseline.

Appendix A

Table A1. Data Collection – April 2014-Dec 2016 June 2015

		April 2014-Dec 2014	
Grades 4, 5	All (N = 3,026)	*Treatment* (N = 2,016)	*Control* (N = 1,010)
April 2014 EGRA/EGMA scores	100%	100%	100%
Oct 2014 Assigned Home Visit 1	69%	84%	40%
Oct 2014 Received Home Visit 1	36%	43%	20%
Dec 2014 EGRA/EGMA scores	89%	89%	90%
		June 2015-June 2017	
Grades 3, 4, 5	All (N = 4,371)	*Treatment* (N = 2,912)	*Control* (N = 1,459)
June 2015 EGRA/EGMA scores: Grade 4, 5 in 2014	86%	85%	87%
June 2015 EGRA/EGMA scores: Grade 3 in 2014	100%	100%	100%
Oct 2015 Assigned Home Visit 2	100%	100%	100%
Oct 2015 Received Home Visit 2	67%	69%	64%
Dec 2015 EGRA/EGMA scores	69%	68%	69%
June 2016 EGRA/EGMA scores	87%	88%	86%
Dec 2016 EGRA/EGMA scores	84%	84%	85%
June 2017 EGRA/EGMA scores	77%	77%	77%

Notes: June 2016 and Dec 2016 EGRA/EGMA assessments included only those students who were enrolled in either grade 3 or grade 4 in 2014. Students who were in grade 5 in 2014 were not followed in 2016. June 2017 EGRA/EGMA assessments included only those students who were enrolled in grade 3 in 2014. Students who were in grades 4 or 5 in 2014 were not followed in 2017.

Table A2. Baseline balance table

	Analysis sample	Treatment	Control	Difference
Baseline test scores				
Words read	91.91	91.68	92.38	-0.69
	(20.09)	(20.18)	(19.91)	[0.55]
Number of correct subtractions	12.75	12.67	12.90	-0.23
	(5.29)	(5.24)	(5.39)	[0.19]
Number of correct sums	17.24	17.24	17.23	0.01
	(5.23)	(5.25)	(5.19)	[0.27]
Number of correct problems	3.81	3.81	3.82	-0.00
	(1.64)	(1.65)	(1.62)	[0.04]
Child characteristics				
Age	9.90	9.91	9.88	0.04
	(1.52)	(1.54)	(1.48)	[0.05]
Gender - female	0.46	0.45	0.47	-0.02
	(0.50)	(0.50)	(0.50)	[0.01]
Grade	4.06	4.06	4.06	-0.00
	(0.82)	(0.82)	(0.82)	[0.00]
Observations	4,371	2,912	1,459	

First and second columns contain means with standard deviations in parentheses. For students in grades 4 and 5 in 2014, the baseline test score was collected in April 2014. For students in grade 3 in 2014, the baseline test score was collected in June 2015, when students were in grade 4. Observations in the first column include all students who were ever assigned to the treatment condition. This includes students who received the treatment in December 2014 and 2015, and students who received the treatment in December 15. The third column includes the difference between students in the treatment and control groups, with asterisks indicating the p-value from a regression of the row variable on an indicator treatment status and grade indicators. Standard error in brackets. Standard error clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p<0.01$.

Table A3. Correlations between measures of student performance over time and across subjects

			Grade 5 in 2014			
	Baseline	Dec 2014	June 2015	--	--	Between math and reading
Baseline	1					0.135*
Dec 2014	0.557*	1				0.223*
June 2015	0.519*	0.561*	1			0.231*
Dec 2015	0.422*	0.448*	0.513*			0.251*
			Grade 4 in 2014			
	Baseline	Dec 2014	June 2015	Dec 2015	June 2016	
Baseline	1					0.062+
Dec 2014	0.459*	1				0.187*
June 2015	0.432*	0.528*	1			0.254*
Dec 2015	0.403*	0.472*	0.618*	1		0.240*
June 2016	0.400*	0.425*	0.486*	0.491*	1	0.343*
Dec 2016	0.324*	0.430*	0.531*	0.497*	0.508*	0.357*
			Grade 3 in 2014			
	Baseline	Dec 2015	June 2016	Dec 2016	--	
Baseline	1	1				0.197*
Dec 2015	0.459*	1				0.194*
June 2016	0.432*	0.528*	1			0.289*
Dec 2016	0.403*	0.472*	0.618*	1		0.280*
June 2017	0.400*	0.425*	0.486*	0.491*		0.362*

+p<0.05 *p<0.01

Table A4. Differential test-taking and missingness

	Missing outcome Dec 2014	Missing outcome June 2015	Missing outcome Dec 2015	Missing outcome June 2016	Missing outcome Dec 2016	Missing outcome June 2017
Treatment	0.004	0.025	0.004	-0.020	0.000	-0.006
	(0.011)	(0.016)	(0.015)	(0.015)	(0.016)	(0.020)
Constant	0.104	0.127	0.312	0.139	0.155	0.234
N	3,026	3,026	4,371	2,765	2,765	1,345
Overall % missing	11%	14%	31%	13%	16%	23%

Estimates from regression of indicator for whether student was missing test score in the follow-up wave on an indicator for treatment status. Standard errors clustered by school at baseline. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5. Initial Home Visit Balance Table

Variable	Analysis sample	Treatment	Control	Difference
Age of the guardian	38.69	38.63	38.89	-0.37
	(10.54)	(10.51)	(10.64)	[0.48]
Working - father/stepfather	0.88	0.88	0.87	0.01
	(0.32)	(0.32)	(0.33)	[0.02]
Working - mother/stepmother	0.46	0.45	0.48	-0.02
	(0.50)	(0.50)	(0.50)	[0.02]
Years of education - father/stepfather	7.63	7.58	7.77	-0.18
	(3.77)	(3.77)	(3.75)	[0.21]
Years of education - mother/stepmother	8.27	8.29	8.21	0.13
	(3.63)	(3.61)	(3.68)	[0.17]
Family income (in # of minimum salaries)	2.04	2.04	2.05	0.00
	(0.80)	(0.79)	(0.82)	[0.04]
Involvement - Ask about school (days/week)	4.85	4.86	4.80	0.07*
	(0.57)	(0.53)	(0.69)	[0.03]
Involvement - Help studying (days/week)	4.15	4.18	4.06	0.15**
	(1.41)	(1.39)	(1.46)	[0.07]
Involvement - Read with her/him (days/week)	3.14	3.16	3.07	0.12
	(1.89)	(1.87)	(1.93)	[0.10]
Involvement – Help with homework (days/week)	4.60	4.61	4.58	0.04
	(0.97)	(0.98)	(0.96)	[0.03]
Involvement – Ask about grades (days/week)	4.64	4.65	4.61	0.03
	(0.93)	(0.91)	(1.00)	[0.06]
Observations	2,057	1,558	499	

First and second columns contain means with standard deviations in parentheses. Observations in the first column include all students who were ever assigned to the treatment condition, and received an initial home visit. This includes students who received the treatment in Dec 2014 and Dec 2015, and students who received the treatment in Dec 15. The third column includes the difference between students in the treatment and control groups, with asterisks indicating the p-value from a regression of the row variable on an indicator treatment status and grade indicators. Standard error in brackets. Standard error clustered at the school level. * p < 0.10, ** p < 0.05, *** p<0.01.

Table A6. Impact on parent behavior

	(1) Parent index	(2) Parent index
Any treatment	0.088	0.054
	(0.055)	(0.069)
Treatment * Low baseline math and/or reading performance		0.084
		(0.112)
Impact on low-performing students		0.139
Observations	1869	1869

Outcome is calculated by the following: Individual measures of parent behaviors (days/week asked about school, helped with studying, read with child, helped with homework, asked about grades) were standardized with respect to the overall sample. The average of these standardize measures was standardized again to form the composite index. All models include controls for age, gender, and baseline composite test score. Standard errors in parentheses. Standard errors clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7. Separately by grade – Using only students with outcome information in all grades

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Grade 5 in 2014						
Treatment	0.014	0.088	0.124*	--	--	--
	(0.063)	(0.053)	(0.062)			
Observations	894	894	894			
Grade 4 in 2014						
Treatment	-0.004	-0.104	0.047	0.260*	0.040	--
	(0.067)	(0.099)	(0.076)	(0.142)	(0.120)	
Observations	780	780	780	780	780	
Grade 3 in 2014						
Treatment	--	--	0.024	0.077	-0.059	-0.144
			(0.068)	(0.074)	(0.083)	(0.113)
Observations			787	787	787	787

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated for baseline and each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8. Impact of treatment on being at score ceiling – Separately by subject

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Student at ceiling for number of words read						
Treatment	-0.000	0.050***	0.034*	0.005	-0.003	0.007
	(0.020)	(0.017)	(0.019)	(0.023)	(0.011)	(0.011)
Observations	2689	2587	2990	2653	2527	1036
Control mean	0.37	0.28	0.44	0.26	0.06	0.04
Student at ceiling for number of correct subtractions						
Treatment	-0.019***	-0.013	0.005	0.012	-0.003	-0.011
	(0.006)	(0.011)	(0.010)	(0.009)	(0.007)	(0.012)
Observations	2522	2449	2898	2596	2475	1036
Control mean	0.07	0.04	0.08	0.04	0.06	0.04

All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9. Impact on student reading outcomes, by gap between baseline reading student performance and parent beliefs

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Treatment	-0.069	-0.069	0.041	0.365***	0.029	0.087
	(0.071)	(0.087)	(0.049)	(0.109)	(0.117)	(0.162)
Treatment* Abs(Gap, reading)	0.002	0.008	0.005	0.019***	0.011*	-0.006
	(0.003)	(0.005)	(0.003)	(0.007)	(0.006)	(0.009)
Observations	985	949	1513	1364	1312	764
Control mean	0.021	1.290	1.245	3.309	2.709	2.655

Reading scores were standardized within grade with respect to the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10. Impact on student subtractions outcomes, by gap between baseline reading student performance and parent beliefs

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Treatment	-0.142	0.018	0.024	0.085	-0.041	-0.063
	(0.100)	(0.054)	(0.051)	(0.052)	(0.068)	(0.081)
Treatment* Abs(Gap, subtractions)	0.021	-0.005	-0.017	-0.025*	-0.019	-0.006
	(0.015)	(0.010)	(0.012)	(0.014)	(0.014)	(0.015)
Observations	939	910	1484	1361	1309	762
Control mean	0.662	0.609	0.733	0.759	0.901	0.761

Subtractions scores were standardized within grade with respect to the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11. Impact on composite test score outcomes, by gender

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Treatment	-0.095**	-0.035	0.003	0.102	-0.207*	-0.034
	(0.044)	(0.073)	(0.043)	(0.098)	(0.114)	(0.142)
Treatment*Gender - Female	0.128	0.096	0.171**	0.010	0.277**	0.028
	(0.089)	(0.109)	(0.069)	(0.116)	(0.106)	(0.203)
Observations	2661	2593	2994	2418	2336	1036
Control mean	0.451	1.288	1.334	2.740	2.422	2.208

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group at baseline. The sum of the standardized reading and math scores were calculated for baseline and each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12. Impact on composite test score outcomes, by baseline performance

	Dec. 2014	June 2015	Dec 2015	June 2016	Dec 2016	June 2017
Low baseline performance: Below 25th percentile in math or reading at baseline						
Treatment	-0.121***	-0.053	-0.017	0.013	-0.081	0.001
	(0.039)	(0.047)	(0.040)	(0.079)	(0.087)	(0.126)
Treatment*Low baseline perf.	0.222***	0.134	0.206**	0.255	-0.011	-0.051
	(0.065)	(0.096)	(0.079)	(0.161)	(0.151)	(0.229)
Impact on low-performing students	0.101**	0.081	0.189***	0.268**	-0.092	-0.050
Observations	2510	2445	2894	2367	2290	1036
Measure of baseline performance: Continuous baseline composite math and reading score						
Treatment	-0.033	-0.005	0.068*	0.122**	-0.090	-0.020
	(0.025)	(0.046)	(0.036)	(0.060)	(0.100)	(0.113)
Treatment*Baseline composite score	-0.051*	-0.049	-0.046	-0.018	-0.002	-0.013
	(0.029)	(0.032)	(0.034)	(0.061)	(0.062)	(0.104)
Observations	2510	2445	2894	2367	2290	1036
Low baseline performance: Below 25th percentile on composite math and reading score						
Treatment	-0.081**	-0.039	0.020	0.085	-0.080	-0.033
	(0.037)	(0.045)	(0.040)	(0.085)	(0.092)	(0.116)
Treatment*Low baseline perf.	0.202*	0.148	0.207*	0.147	-0.028	0.051
	(0.102)	(0.099)	(0.115)	(0.251)	(0.199)	(0.259)
Impact on low-performing students	0.121	0.109	0.227**	0.232	-0.108	0.019
Observations	2510	2445	2894	2367	2290	1036
Low baseline performance: Below 50th percentile on composite math and reading score						
Treatment	-0.092**	-0.026	0.001	0.054	-0.059	-0.033
	(0.045)	(0.062)	(0.045)	(0.099)	(0.096)	(0.137)
Treatment*Low baseline perf.	0.121*	0.047	0.134	0.131	-0.063	0.027
	(0.067)	(0.081)	(0.091)	(0.163)	(0.163)	(0.195)
Impact on low-performing students	0.029	0.021	0.136*	0.184*	-0.122	-0.006
Observations	2510	2445	2894	2367	2290	1036

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated, and standardized again with respect to the control group, to form the composite score. All models include controls for age, gender, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13. Impact on math and reading outcomes, by baseline performance

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Number of correct words read (standardized)						
Treatment	-0.024	-0.034	-0.016	-0.008	-0.122	-0.045
	(0.036)	(0.041)	(0.027)	(0.092)	(0.101)	(0.142)
Treatment*Low baseline reading	0.179**	0.227	0.299**	0.289	0.116	0.241
	(0.078)	(0.150)	(0.119)	(0.286)	(0.291)	(0.372)
Impact on low-performing students	0.156*	0.193	0.283**	0.281	-0.006	0.196
Observations	2689	2587	2991	2414	2332	1036
Number of correct subtractions (standardized)						
Treatment	-0.147***	-0.031	0.035	0.107***	-0.037	-0.013
	(0.053)	(0.043)	(0.040)	(0.039)	(0.044)	(0.082)
Treatment*Low baseline math	0.286***	0.022	0.064	-0.035	0.011	-0.144
	(0.093)	(0.088)	(0.068)	(0.078)	(0.084)	(0.120)
Impact on low-performing students	0.139**	-0.009	0.099	0.072	-0.026	-0.157
Observations	2522	2449	2898	2371	2294	1036

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated, and standardized again with respect to the control group, to form the composite score. All models include controls for age, gender, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14. Impact on math and reading outcomes, by baseline performance

	(1) Dec. 2014	(2) June 2015	(3) Dec 2015	(4) June 2016	(5) Dec 2016	(6) June 2017
Number of correct words read (standardized)						
Treatment	-0.027	-0.013	-0.034	-0.054	-0.082	0.044
	(0.039)	(0.050)	(0.032)	(0.095)	(0.106)	(0.149)
Treatment*Low baseline reading or math	0.135**	0.106	0.198**	0.326	-0.031	-0.054
	(0.060)	(0.125)	(0.082)	(0.214)	(0.196)	(0.276)
Impact on low-performing students	0.107*	0.093	0.164**	0.273*	-0.114	-0.010
Observations	2541	2445	2897	2367	2290	1036
Number of correct subtractions (standardized)						
Treatment	-0.152***	-0.064*	0.012	0.082*	-0.033	-0.042
	(0.050)	(0.038)	(0.047)	(0.048)	(0.050)	(0.089)
Treatment* Low baseline reading or math	0.192***	0.091*	0.097	0.036	-0.001	-0.026
	(0.065)	(0.050)	(0.064)	(0.067)	(0.075)	(0.153)
Impact on low-performing students	0.040	0.027	0.109**	0.118**	-0.034	-0.067
Observations	2517	2445	2894	2367	2290	1036

Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated, and standardized again with respect to the control group, to form the composite score. All models include controls for age, gender, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

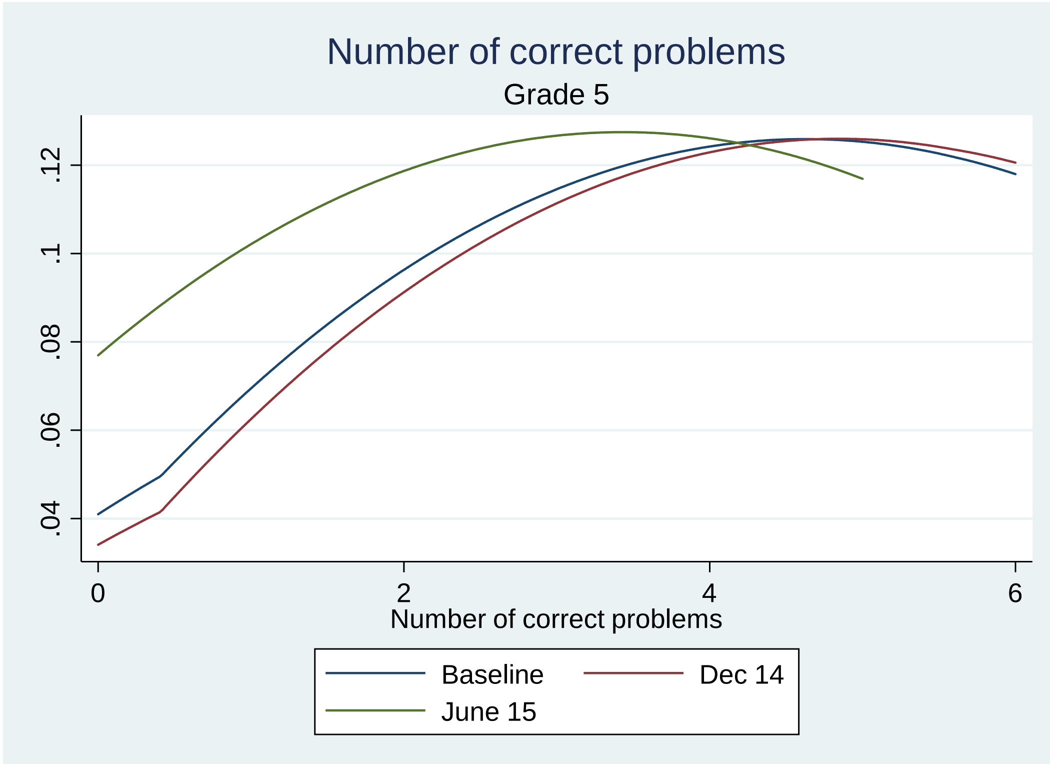

Figure A1. Distribution of number of correct problems across first three study waves

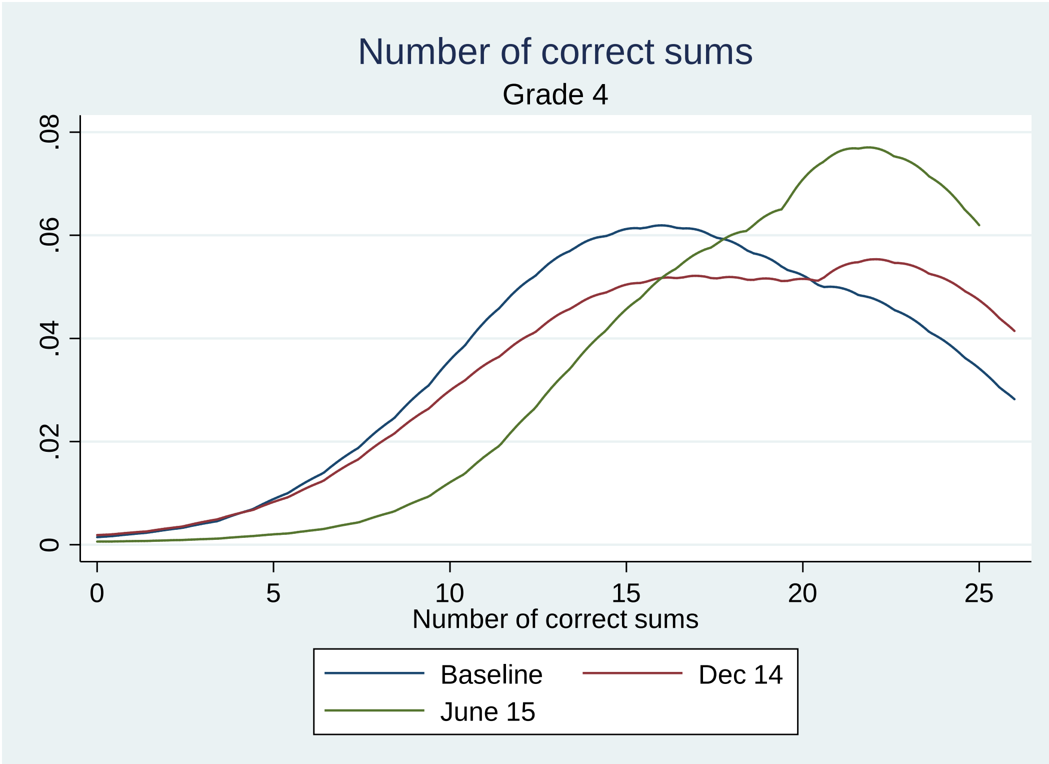

Figure A2. Distribution of number of correct sums across first three study waves

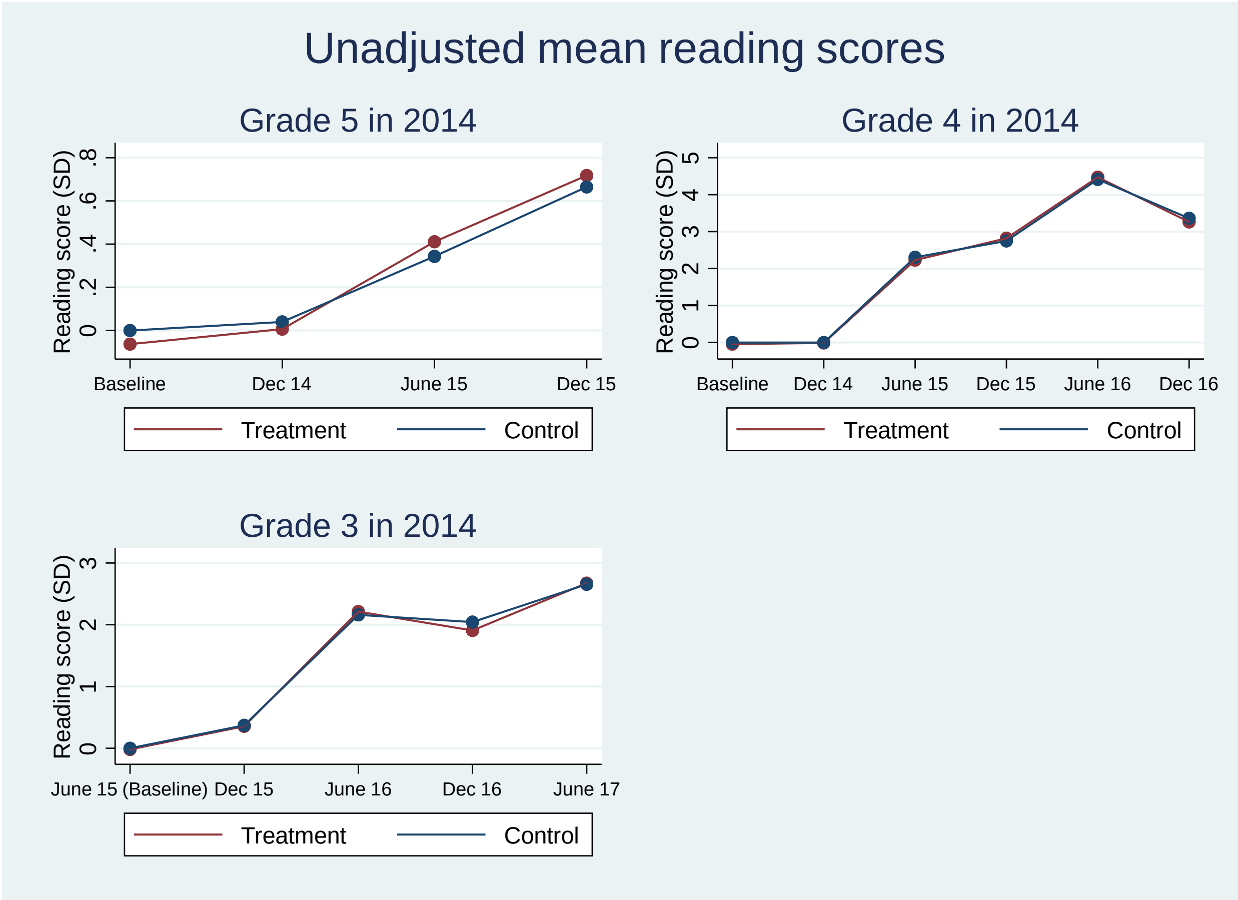

Figure A3. Unadjusted standardized reading scores across follow-up waves. Note: Reading scores standardized with respect to the control group at baseline.

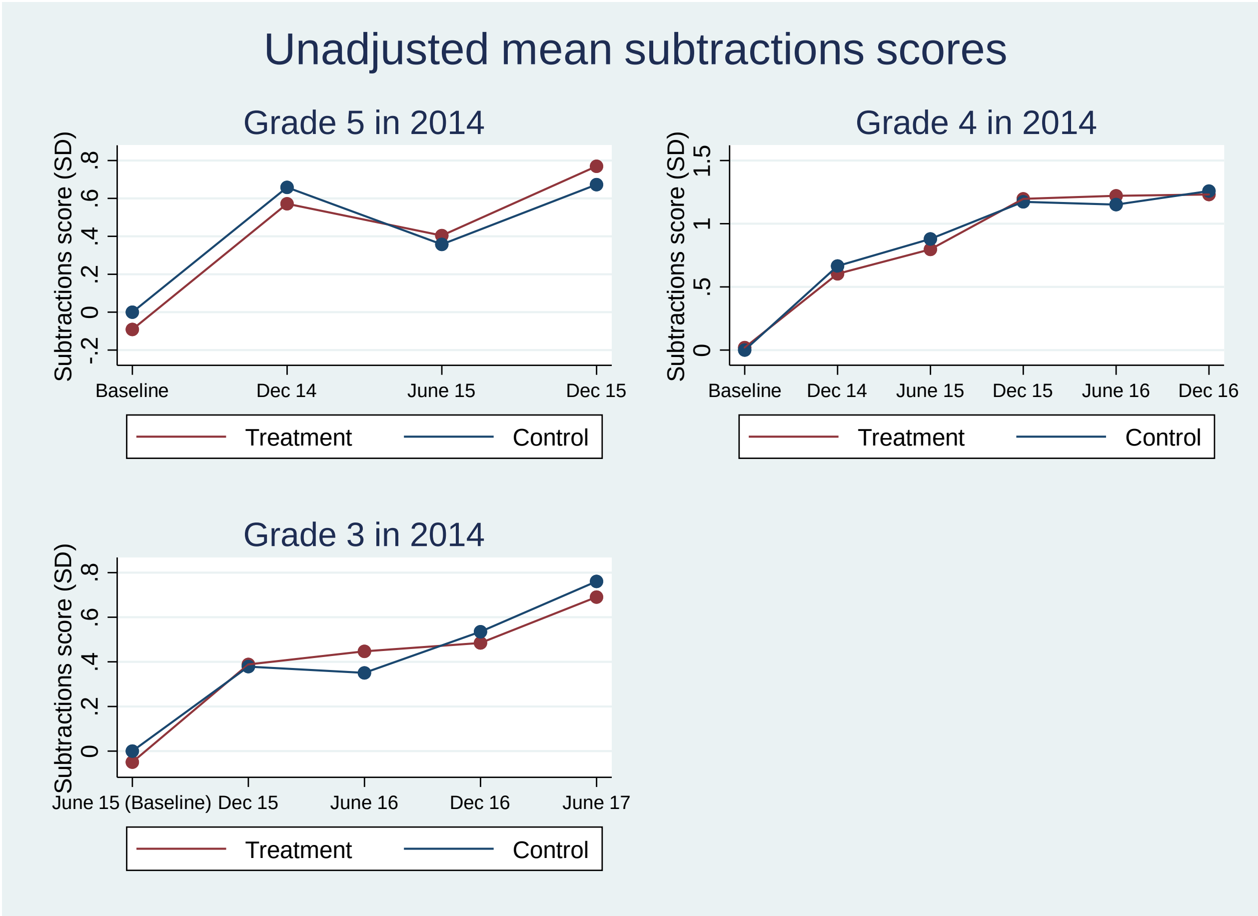

Figure A4. Unadjusted standardized subtractions scores across follow-up waves. Note: Subtractions scores standardized with respect to the control group standard at baseline.

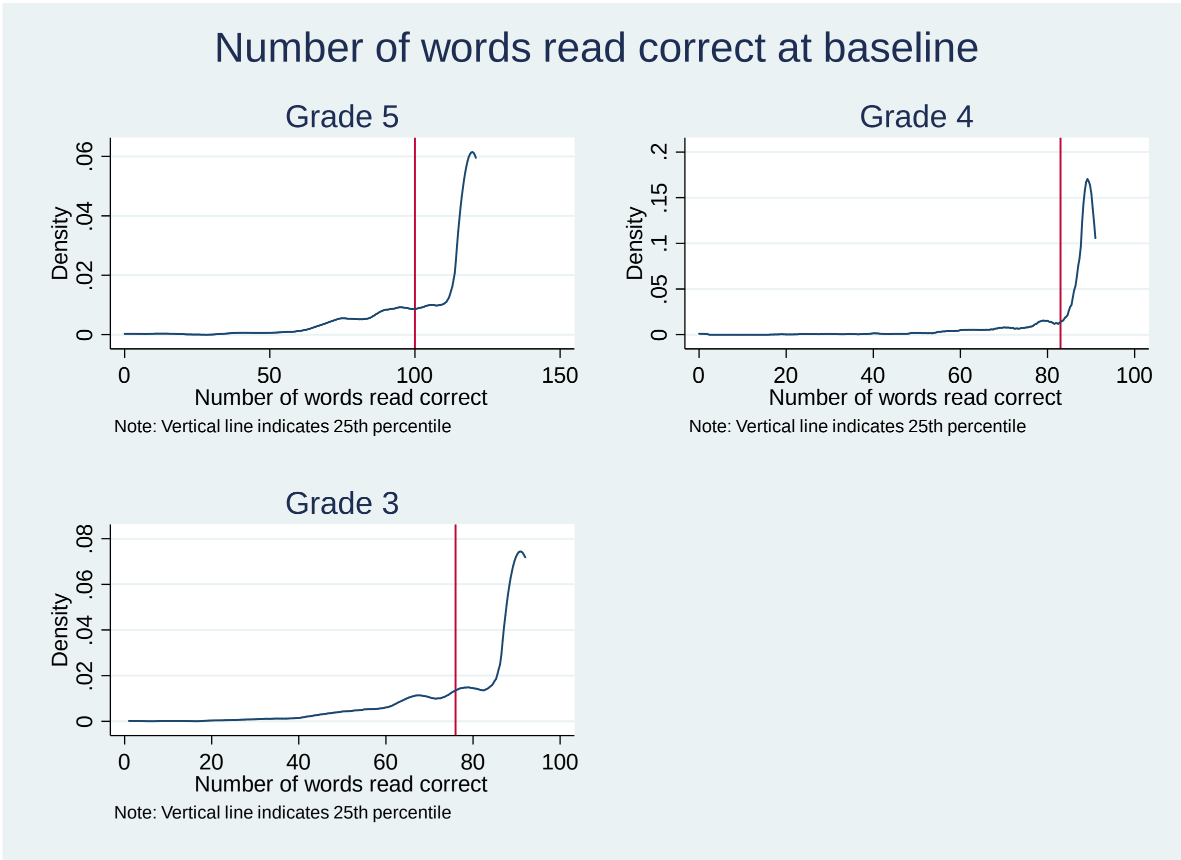

Figure A5. Baseline reading scores: above vs. below 25th percentile. Note: Vertical lines indicate scores at 25th percentile

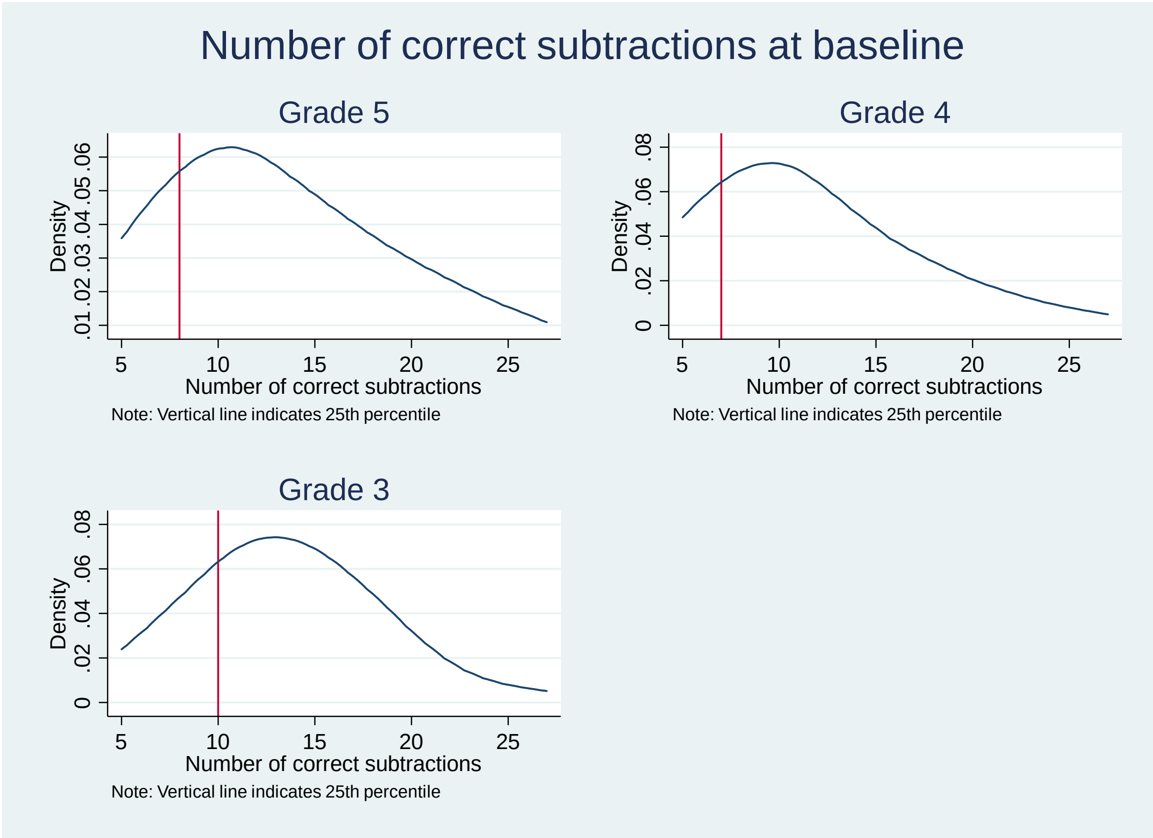

Figure A5. Baseline reading scores: above vs. below 25th percentile. Note: Vertical lines indicate scores at 25th percentile

Appendix B

Details and results of the family-engagement intervention

Description of the family-engagement intervention

In the third phase of the study (2016-2017) we incorporated a family-engagement component. This component focused on the teachers of the students already in the experiment. In July 2016, teachers were randomly assigned to one of two groups. In the treatment group (Family-engagement intervention), teachers received a report card containing test score information, collected in June 2016, for the students in their class. We provided teachers with information on all students in their class, regardless of whether they had received they had participated in Phase 1 or Phase 2 of the study, and regardless of whether the household had received individual information or not. Teachers also received a list of suggestions to promote family-school engagement, with two components. First, a list of suggestions on how to improve their communication with the families. Second, a list of suggestions on how to encourage families to engage with their children's education outside of the school. No information was provided to teachers in the control group.

To implement this intervention, teachers were visited at their schools. A questionnaire was administered to all teachers. Only teachers assigned to the treatment group received the report card mentioned above. In December 2016 and June 2017, we administered new rounds of EGRA and EGMA tests to all students in our sample.

Results of the family-engagement intervention

We test two specifications to examine the impact of the teacher intervention. In the first specification, we examine the differences in students' results based on whether they were in a classroom with a treated or control teacher. In the second specification, we test the interaction between teacher group (treatment/control) and students group (treatment/control). The effects of the intervention of teachers are null.

It may be the case that the intervention did not provide new information to teachers; teachers had knowledge of their students' abilities but were unable to act upon it. Alternatively, the intervention may have provided new information to teachers, but the receipt of this new information did not change teacher behaviors. did in fact provide new information to teachers, but teachers were unable to act on it. We can not rule out any of these (plausible) hypotheses.

Table B1. Impact of the family-engagement intervention

	Dec. 2016	Dec. 2016	June 2017	June 2017
Math and reading composite				
Teacher treatment	-0.073	-0.017	-0.174	0.022
	(0.100)	(0.130)	(0.135)	(0.209)
Teacher treatment* Individual treatment		-0.083		-0.290
		(0.141)		(0.197)
Observations	2049	2049	965	965
Control mean	2.260		2.286	

Teacher treatment = Teacher was assigned to the treatment condition of the family-engagement intervention in Phase 3. Individual treatment = Student was assigned to the treatment condition of the household information intervention in Phase 1 or Phase 2. Composite math and reading score calculated by the following: Reading and math (subtractions) scores were standardized within grade with respect to the control group. The sum of the standardized reading and math scores were calculated for baseline and each wave. Composite scores for each follow-up wave were standardized again with respect to the composite score for the control group at baseline. All models include controls for age, gender, grade, and baseline math and reading scores. Standard errors in parentheses and clustered at the school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.