NBER WORKING PAPER SERIES

REDUCING PARENT-SCHOOL INFORMATION GAPS AND IMPROVING EDUCATION OUTCOMES: EVIDENCE FROM HIGH-FREQUENCY TEXT MESSAGES

Samuel Berlinski
Matias Busso
Taryn Dinkelman
Claudia Martínez A.

Working Paper 28581
http://www.nber.org/papers/w28581

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021, Revised June 2022

Julian Martinez-Correa provided excellent research assistance and multiple comments that improved the manuscript. Anna Koh Lee, Santiago Perez Vincent, Dario Romero, and Dario Salcedo provided excellent research assistance in early stages of the paper. We thank Bernardita Muñoz, Daniela Alvarado and Paula Espinoza for superb support in fieldwork. Thomas Dishion and Anne Mauricio gave fabulous guidance on communicating with parents and the Family Check-Up approach. The authors gratefully acknowledge funding through J-PAL's Post-Primary Education Initiative, the Inter-American Development Bank, the Spencer Foundation, and a Chilean FONIDE grant (No. 711272). IRB approval: MIT COUHES Protocol # 1308005856. The study is registered as AEARCTR-0000458. 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Reducing Parent-School Information Gaps and Improving Education Outcomes: Evidence from High-Frequency Text Messages
Samuel Berlinski, Matias Busso, Taryn Dinkelman, and Claudia Martínez A.
NBER Working Paper No. 28581
March 2021, Revised June 2022
JEL No. D8,I25,N36

ABSTRACT

We conducted an experiment in low-income urban schools in Chile to test the effects and behavioral changes triggered by a program that sends attendance, grade, and classroom behavior information to parents via weekly and monthly text messages. Our 18-month intervention raised average math scores by 0.09 of a standard deviation and increased the share of students satisfying attendance requirements for grade promotion by 4.7 percentage points. Treatment effects were larger for students at higher risk of later grade retention and dropout. Our results show that leveraging existing school inputs in a light-touch, cost-effective, and scalable information intervention can improve school outcomes in low-capacity settings.

Samuel Berlinski
Research Department
Inter-American Development Bank
1300 New York Avenue, NW
Washington, DC 20577
samuelb@iadb.org

Matias Busso
Research Department
Inter-American Development Bank
1300 New York Ave. NW
Washington, DC
mbusso@iadb.org

Taryn Dinkelman
Department of Economics
3058 Jenkins Nanovic Halls
University of Notre Dame
South Bend, IN 46617
and CEPR
and also NBER
tdinkelm@nd.edu

Claudia Martínez A.
Instituto de Economia
Pontificia Universidad Catolica de Chile
Av. Vicuna Mackenna 4860
Macul, Santiago
Chile
clmartineza@uc.cl

A data appendix is available at http://www.nber.org/data-appendix/w28581

1 Introduction

Grade retention and early dropout are two of the biggest challenges facing education systems in many middle-income countries today. In Latin America, only 46% of students graduate from secondary school on time, and only 53% of young people aged 20 to 24 have completed their high school education (Busso et al. 2017). These poor schooling outcomes contribute to persistent education gaps between low- and high-income families.

Researchers have identified absenteeism, failing grades, and classroom misbehavior as important early warning signals for grade retention and the likelihood that students will eventually drop out of school (Manacorda 2012, Wedenoja 2017). While schools around the world routinely record these types of student outcomes, families often do not have timely access to this information. In this paper, we examine whether increasing the frequency and ease of communication between parents and schools can improve students' academic outcomes, particularly among those who are at higher risk of being retained at a given grade or of later dropout. We evaluate an intervention that leverages existing school resources and practices to improve education outcomes. We also explore several channels through which this intervention may have changed parenting practices around schooling.

In 2014 and 2015, we conducted a randomized experiment in Chile to evaluate the effects of using weekly and monthly cellphone text messages to provide parents with up-to-date information on students' attendance, grades, and classroom behavior. The intervention focuses on students in the last five grades of primary school, years during which attendance and grades start to matter, but before the risks of grade repetition or dropout significantly increase. The text message intervention (Papas al Dia) was deliberately designed to be a low-touch intervention, with no change in behavior required by schools or teachers. Importantly, we sustained the high-frequency text messaging over two school years, to allow parents time to adapt their parenting strategies in response to an ongoing flow of student-level information.

Our main experimental sample includes about 1,000 children enrolled in seven low-income schools in a metropolitan area in Chile. After conducting baseline student and parent surveys and collecting school administrative data on student outcomes, we randomly varied which classrooms in each school were to receive a high (75%) or low (25%) share of treated students, and then randomized individual students in each classroom into the text messages treatment. Over 18 months, we delivered more than 44,000 text messages to families in our sample. Treatment messages containing information about attendance, grades, and behavior were sent to treated parents, while control parents received general all-school text messages during this time. We continued to collect administrative data throughout the two years and conducted mid- and endline parent and student surveys. Our data allow us to measure schooling outcomes as well as changes in parent information sets and parenting practices.

We begin by documenting sizable gaps that exist between parents' knowledge and school reports of students' attendance and grades. Comparing baseline survey responses to school records, we find that 26 percent of parents were unable to report correct information about their child's grades; while 48 percent could not approximate their child's school attendance in the previous two weeks. Similar information gaps have been found in settings as diverse as the United States (Bergman 2021), Malawi (Dizon-Ross 2019) and Colombia (Barrera-Osorio et al. 2020). Moreover, we document that the parents of at-risk, low-achieving students are more likely to misreport grades and attendance at baseline. Narrowing this gap –between parents' understanding of their child's performance and actual performance as documented by the school– is a key target of our text messaging treatment. Parents who have more accurate knowledge about recent grades, attendance, and behaviors are likely to be more engaged with their child's schooling on a day-to-day basis in ways that improve schooling outcomes (Escueta et al. (2020), JPAL (2020)).

Our main results are that exposure to the messaging treatment had positive impacts on math grades and attendance, with particularly large impacts on at-risk students, and positive spillover effects within classrooms. Relative to control students, treated students increased their math GPA by 0.09 of a standard deviation, and the probability of treated students earning a passing grade in math increased by 2.7 percentage points (or 2.9% relative the control mean of 93%). The intervention increased school attendance by 1.1 percentage points (or 1.2% relative to the control mean of 87%), and increased the share of students who satisfied the attendance requirements for grade promotion by 4.7 percentage points (or 6.4%, relative to a control mean of 73%). On average, there were no significant impacts of the treatment on recorded misbehavior in school. We find important heterogeneity in these treatment effects related to initial academic performance. Grades and attendance impacts are 40-60% larger, and misbehavior falls by a significant 0.2 standard deviations more, among those students with one standard deviation more of our at-risk index.

Exploiting aspects of the research design and using our detailed administrative data, we investigate some of the ways in which the information intervention operated on parents and students. First, using variation in the weekly and monthly frequency of text messages delivered, we examine whether the effects of messages changed over time or with the frequency of the messaging. The patterns in our data indicate that the positive effect on attendance fades out over the week: effects appear somewhat larger immediately after parents receive the text messages and decline as the days go by. This suggests that for outcomes where the student makes daily choices –to attend or not to attend school– high-frequency text messages may be more beneficial than sporadic messages. At the same time, we find that the intervention is effective throughout the school year. Despite the sustained nature of the treatment, parents do not seem to "get used" to the treatment. Although the data do not allow us to precisely estimate all of the patterns of effects related to timing and frequency of messaging, taken together, the results suggest that information treatments like the one studied in this paper may be more effective when delivered at high frequency and in an ongoing way over time.

Next, we use the random manipulation of the share of treated students in each classroom to assess spillover effects within treated students. Understanding spillovers is important for thinking about impacts when information interventions like these scale. Although our design does not allow us to test for spillovers to the control group, we find evidence of positive classroom-level spillovers among treated students. This suggests that the positive direct effect on individual grades and attendance that we measure likely underestimates the impacts of a scaled-up version of this program in which all students would be treated.[1]

The information intervention was targeted at improving communication between parents and schools, thereby lowering parent monitoring costs and enabling better parent engagement with students and with schools. We use our rich administrative data and information collected through surveys conducted with parents and students before and after the program to explore these channels. We show that exposure to the high-frequency text message treatment shrinks information gaps about math scores and misbehavior between parents and schools. Parents of at-risk students "correct" their understanding of their child's performance to the greatest degree (although results are not statistically significant at conventional levels). And, although the information treatment was designed to deliver information about specific subjects and behaviors, we show that it may also have directed parents to pay more attention to all aspects of school performance: the treatment group performed better in non-targeted subjects (e.g., language), and parent misinformation about these non-targeted subjects also improved among the treated group.

Suggestive evidence from our surveys (point estimates are not always statistically significant) indicates that treated parents used the new information they obtained about their children to guide interactions with their children at home. Treated students report significantly more family support as a result of the intervention and that their parents were more involved in school matters. Parent engagement in day-to-day school matters appears to have changed as a result of the sustained, high-frequency information intervention. Consistent with these changes in reported parental behavior, we find that a large share of parents are willing to pay for the information program. We rely on a survey experiment to assess willingness to pay for access to the information program. For all parents, demand slopes downward:

[1] For budget reasons we do not have pure control classrooms, therefore we are restricted to estimating spillovers within treated students.

over 70% of parents are willing to pay for the text messaging service when offered the lowest randomized price and this share falls as the randomized price rises.[2] We cannot reject that treated parents have the same elasticity of demand for the program as control parents.

Our study contributes to a large and active literature studying the effect of sending parents information about their children's activities and performance in school. This literature includes Bursztyn & Coffman (2012), Kraft & Dougherty (2013), Avvisati et al. (2014), Castleman & Page (2015), Kraft & Rogers (2015), De Walque & Valente (2018), Rogers & Feller (2018), Bergman & Chan (2021), Dizon-Ross (2019), Angrist, Bergman & Matsheng (2020), Barrera-Osorio et al. (2020), Bergman (2021), Bergman et al. (2020), Gallego et al. (2020), and Bettinger et al. (2021) among others.[3] Some of the key findings in this literature show that bridging information and communications gaps between parents and schools, however it happens (by text messages, email, regular phone calls, regular mail, report cards, or in-person visits), often results in learning gains for students.

We make three main contributions to this literature. Our study is the first to implement an information treatment sustained for almost two school years. The unusually long duration of our intervention (18 months) contrasts with prior studies evaluating once-off information interventions, or interventions that last several months. The duration of an information treatment may matter for several reasons. Continuing the text message program over multiple years means that parents experience a persistent, rather than one-time, improvement in information and reduction in monitoring costs. In the face of this persistent change in parent-school communications, parents may have been able to adopt different types of parenting strategies than they otherwise would have after a one-time or shorter-lived treatment (e.g. engaging more with schools, or providing more family support for schoolwork, as students here report). In addition, the value of some types of information (e.g. attendance this week, or grade on a recent test) likely falls over time: for example, parents may be most likely to act on truancy in the days or weeks following a reported event. A sustained information treatment allows parents to always be up-to-date with this type of information. Furthermore, since the novelty of receiving information might fade-out over time, it is important to test for ongoing impacts in a long-term treatment. Overall, we interpret the results from our sustained treatment as providing a good sense of how parents would respond and how student outcomes would change, on average, in a realistic environment outside of an experiment.

Our second contribution is to test an intervention that uses primarily *existing* school

[2] This result echoes the findings of Bursztyn & Coffman (2012), who show that Brazilian parents report being willing to pay for receiving regular updates on their child's absenteeism.

[3] Online Appendix A provides a summary of the results in these papers.

inputs, expanding the potential for cost-effective scalability in low-capacity school settings. Implementing a parent-school communications program like Papas al Dia would entail a low variable cost and a one-time setup cost. In the last section of the paper, we use our main intent-to-treat estimates for grade impacts to show that the program-specific variable cost of achieving a 0.01 of a standard deviation increase in math grades is about US$1.21 per student per year at market prices (rising to US$1.39/year when we include a fixed set up cost for the digital platform). Compared to other interventions in Latin America designed to improve learning outcomes and attendance, a program like Papas al Dia is cost-effective. For example, in a similar setting in Colombia, (Barrera-Osorio et al. 2020) combined a one-time information intervention about student performance in grades 4 to 6 with targeted advice to parents. Their results indicate short-run gains on a combined math and reading test score that are close to the Papas al Dia test score results, but at considerably higher cost per student (US$7.50 per year).

A key feature of our text messaging program contributing to potential scalability is that we did not require any change in teacher inputs, practices, or pedagogy for implementation. Moreover, it was possible for us to implement (and evaluate) the intervention for such a long time because we leveraged existing school practices and high-frequency data already collected by teachers, without making their jobs more complex. In that sense, our paper is most closely related to Bergman & Chan (2021), who automate the process of gathering already-digitized student data (scraping student information systems and feeding this into a text messaging platform) to facilitate an information intervention in 22 schools (covering grades 6-12) in West Virginia.[4] Transitioning teachers to classroom-level digital student information systems like this would require substantial new resources in low- and middle-income countries, where information about students is still often collected on paper. As part of our experiment, we used a small team of research assistants to collect and digitize paper records from each school each week. In practice, a scalable approach could leverage existing administrative school staff to enter information from class books into a digital platform like Papas al Dia, ready for text message dissemination. In this way, programs such as Papas al Dia offer a practical, effective, and low-cost example of how to bridge information gaps between old school paper records and parent cellphones at scale.

A third contribution of our work is to provide evidence on the impacts of an information intervention in a new Chilean setting. As Angrist, Evans, Filmer, Glennerster, Rogers & Sabarwal (2020) note, information interventions to improve learning gains tend to have high variance across settings. Most of the studies mentioned above take place in United States'

[4]Rather than sending regular text messages to all parents, Bergman & Chan (2021) designed their study to only alert parents to missed classes, missed assignments, and low grades.

school systems. Although schools in our study resemble poor schools in United States cities, our treatment effects do differ somewhat from the effects of similar programs implemented in different contexts.[5]

Our estimated learning gains in math (0.09 s.d.) are at the lower end of the range in the literature (0.09-0.19 s.d. of test scores), while our attendance gains (1.1. percentage points) fall in the middle of the range (0-2.1 percentage point gains in attendance).[6] Given the profile of schools in our sample, our intent-to-treat results are likely to be relevant for poor performing schools in urban settings of developed and middle-income developing countries. More generally, an important emerging pattern from our work in Chile, from the work of Barrera-Osorio et al. (2020) in Colombia, and from Bergman & Chan (2021) in the US, is that interventions improving parent-school communications tend to have largest test score effects for the weakest students.[7] Closing information gaps between parents and schools in an effective way, starting as early as late elementary school, may therefore contribute to shrinking achievement gaps in a persistent manner.

2 Setting

There are twelve years of mandatory schooling in Chile: eight of primary school and four of secondary school. Our experiment focuses on children from 4th to 8th grade that attend schools in an urban setting. Children walk to their neighborhood school or take public transportation. Depending on their age, they may travel alone, with an adult, or with older siblings. They attend school for about 180 days in the year, from 8:00 AM to 4:00 PM.[8] After school, children return home and are supposed to do homework. Many are unsupervised when they return home. This set-up is similar to other large urban areas around the world.

Although Chile is now a high-income country, schools still lag behind relative to those in the United States or in the average OECD country. For example, average class size in Chile's secondary schools is 35 students, while in the United States the figure is 26. According to the 2018 PISA results, almost one-third of Chilean students are below the minimum proficiency level in reading compared with 19.3% in the United States; over half of

[5]While not all Chilean schools are low capacity, the schools in our sample have been tagged by the Chilean Ministry of Education as requiring additional resources and support based on poor student outcomes.

[6]These ranges are taken from our summary of the literature in Online Appendix A

[7]This is not the case everywhere. For example, in Malawi Dizon-Ross (2019) finds that better information increases inequality between students as parents are better able to target resources towards the highest ability children. However, schools in Malawi are predominantly rural and have uniformly fewer resources than schools in Chile. The context is very different to ours.

[8]Most schools in Chile have full day schools. Schools can distribute their mandated hours throughout the week, and typically have classes from 8am to 4pm four days a week, ending at 1pm one day a week.

Chilean students (51.9%) are below minimum proficiency in math, compared with 27.1% in the United States. As in many other urban school settings, students are highly segregated into schools by socioeconomic status (Mizala et al. 2007).

Recent high school graduation rates in Chile are around 90%, 10% higher than the average OECD country (OECD 2022). This figure, however, masks considerable inequalities. High school dropout in Chile is concentrated among students in lower-income quintiles. For instance, in 2017, only 79% of students in the lowest-income quintile completed high school, compared with over 96% of students in the highest-income quintile. Attendance, grades, and classroom behavior in elementary school are key factors affecting the risk of grade retention, which, in turn, increases the probability that students will drop out of school when they grow older (e.g. Manacorda 2012, Wedenoja 2017). We focus on these three variables being the early warning signals for poor school outcomes later on.

To advance to the next grade, Chilean students must attend at least 85% of school days in a school year, and obtain a passing grade of 4.0 in all subjects (on a scale from one to seven).[9] As a result, there is a strong correlation between attendance, subject grades, and grade retention.[10]

The transition from the final grade of primary school to the beginning of secondary school is a point at which students are at high risk of grade retention or, in the worst case scenario, of dropping out of the school system. Even though grade retention is an outcome of concern during lower grades, it becomes even more of a concern as students progress through their school years. During grades 1-3 about 3% of students repeat their grade. Starting in grade 4 this percentage increases with each grade, finally reaching 5% by the end of primary school. In the first year of secondary school, the grade retention rate surges, reaching 13%. This pattern is observed in our sample and is common in most Latin American countries (Bassi et al. 2015).

Our intervention focuses on students in the last five grades of primary school, where the median child age is 10. It targets information for parents during the years when attendance, grades, and behavior start to matter, but before the risks of grade repetition or dropout significantly increase.

[9] Students who fail one subject can still advance to the next grade if they maintain an average grade of 4.5 for the remaining subjects; students who fail two subjects can also advance if they maintain an average grade above 5.0 in the remaining subjects. The 85% attendance requirement can be lifted by the school board under special circumstances.

[10] Using administrative data, we examined these same correlations in our sample prior to the start of the intervention. The correlation of average grade was 0.4 with attendance and -0.4 with grade retention. The correlation between school attendance and grade retention was -0.3. Even conditional on age and gender controls, and taking into account grade-level and school fixed effects, the correlations between lower attendance, lower grades, and a higher risk of failing the grade are large and statistically significant at the 5% level.

Gaps in the information that schools and parents have about children have been identified in settings as diverse as the United States (Bergman 2021), Malawi (Dizon-Ross 2019) and Colombia (Barrera-Osorio et al. 2020). Examples in the literature suggest that most parents tend to overestimate their child's performance in school, and that parents who have less education themselves have worse information about their child's performance in school (Barrera-Osorio et al. 2020, Rogers & Feller 2018, Bergman & Chan 2021). Parents in our sample are literate, but have generally low levels of education (e.g. only 53% of mothers have completed high school).

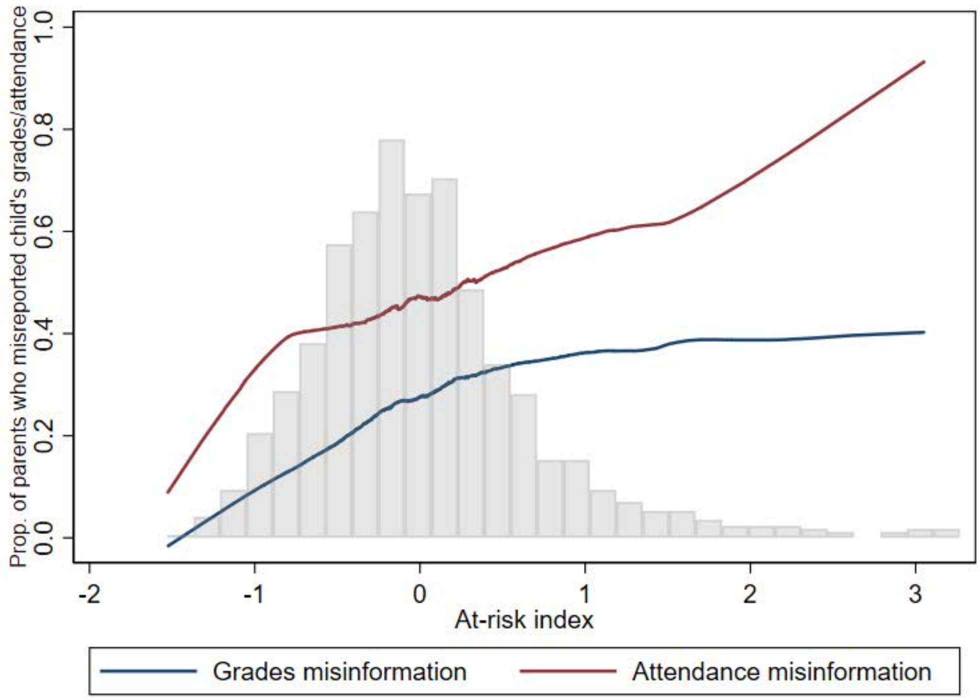
In our setting, we observe similar types of parent-school information gaps regarding the student's actual grades and attendance. Parents are usually provided with information about their child's progress once per quarter through a physical report card that details a student's grades and number of absences. Not all report cards make it home. Teachers and principals also communicate with parents on an "as needed" basis for certain cases of misbehavior, regular absenteeism, and repeated low grades. Figure 1, based on data from our baseline parent survey described in Section 4, plots the share of parents whose report of the child's grade/attendance is at odds with the child's actual school performance before the intervention began. We define a grade as being misreported if it deviates more than 0.5 points above or below the actual grade. The share of grade misreports is plotted in blue. We define attendance to be misreported if the parents' report of the child's absence differs by two or more instances from actual absences recorded in the previous two weeks. The share of attendance misreports is plotted in red.[11] These misreports are graphed against a summary measure – the (standardized) at-risk index – of whether a child is considered at-risk of retention or dropping out (because of higher absenteeism, lower grades, or worse behavior in class) before the intervention.[12] The histogram describes the distribution of this at-risk index.

In our sample, on average, 26 percent of parents were unable to report correct information about their child's grade while 48 percent could not correctly report their child's school attendance in the previous two weeks. Moreover, Figure 1 shows that misinformation is higher among parents of students with higher at-risk index values, and that a larger share of parents misreport attendance, relative to grades, for students at all levels of risk. About 40% percent of parents of students with a baseline math grade below 4.5 did not accurately know their children's test scores. Similarly, 70% percent of parents of students with an attendance rate of lower than 85 percent, did not know how many days their children had missed school

[11]Parents who did not respond to either question were also classified as misinformed. See notes on columns [2] and [4] of Table 5 for details.

[12]We discuss how we construct this at-risk index in Section 4.

Figure 1: Baseline Share of Misinformed Parents



Note: Y-axis presents the (lowess-smoothed) share of parents misinformed regarding their child's grades (blue line) and attendance (red line) for different levels of the at-risk index –whose histogram is shown in grey. Estimates are based on parent surveys and administrative data at baseline. See notes for columns [2] and [4] of Table 5 for details on the construction of misinformation measures and Section 4 for the index construction.

in the previous two weeks. This is despite 79% of parents in our survey declaring that they almost always check their children's report. These are the types of information gaps our intervention is designed to address. The patterns in Figure 1 suggest that our intervention should be particularly relevant for those children who are the most at-risk of grade retention or dropping out.

3 Experimental Design

In this section we outline the basic elements of our experiment: the recruitment of schools and parents, the randomization of students and classrooms, and the intervention.

Recruitment of participants. We recruited publicly-funded schools from across two municipalities in Santiago.[13] Chile's Quality of Education Agency rates schools based on student

[13]There are two types of public schools in Chile: pure public schools and voucher schools. In one municipality we worked with local education officials to recruit public schools. In the second municipality, we recruited a voucher school. Our main sample consists of students in 63 classrooms across seven schools.

learning, student social and personal development, and any recent changes in these measures. The schools in our sample are particularly deprived according to these ratings. Three of our schools (42.9%) are in the "insufficient" category (the lowest category), and two (28.6%) are in the medium and medium-low categories. Nationally, only 7.6% of schools are ranked as "insufficient". Schools in our sample served students of medium-low- or low-socioeconomic status. Learning outcomes in our schools are among the lowest in Chile: in 2015 national standardized tests, our sample schools perform between the 18th and 35th percentiles.

In recruited schools, we held a series of meetings, inviting parents of all students in grade 4 and above to join the experiment.[14,15] Over 50% percent of parents consented to participate. Consent rates by grade-level were similar. Younger students, those not new to the school, and those with better baseline attendance and grades were somewhat more likely to consent.[16]

Randomization and Intervention. We assigned students to treatment in two steps. First, we stratified by school grade-level, and randomly allocated classrooms (sections) to include a high or low share of students whose parents would receive text messages. In high-share classrooms, 75% of students whose parents had consented to participate were treated; in low-share classrooms, 25% of students whose parents had consented were treated.[17] Second, within each classroom, we randomized students whose parents had consented into treatment or control status, according to the shares allocated in the first-step randomization. Students retained their individual and classroom-level randomization status for the duration of the intervention. Teachers were not informed about which students in their classrooms were participating in the experiment, or who was randomized to treatment.[18]

Figure 2 shows the timeline of the intervention and the data collection. The school year in Chile runs from March to December, with two weeks of winter vacation in July. A first welcoming message was sent to all participants in May of 2014. The intervention started

[14]Consent forms were distributed during an initial parent meeting or later sent home with children.

[15]Initially, students whose parents consented to participate in the experiment were in grades 4 to 8 in the eight schools that participated in the study. The composition changed in the second year. Students in grade 8 participated during the first year of the experiment, but these students could not be treated or followed into secondary school. In addition, one school decided not to continue during the second academic year because it chose to allocate internal resources to other school goals. Because randomization was done at the individual level, stratifying by classroom, the main analysis does not include either the school that dropped out of the program, or the students who were in grade 8 at baseline. In the appendix of tables and figures we show the main results when using this "full" sample as a robustness check.

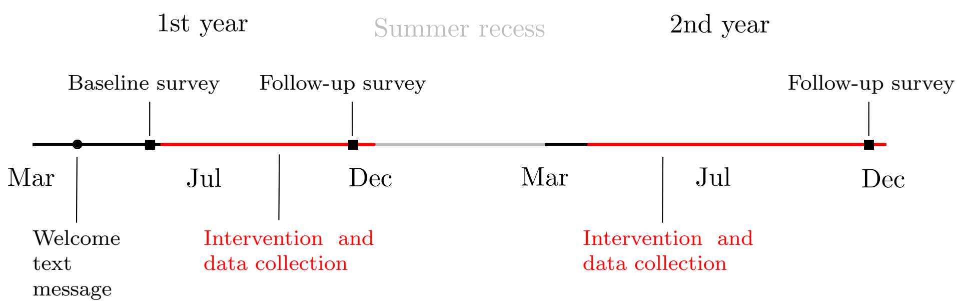
[16]See Online Appendix B for more details regarding the sample and the characteristics of students whose parents consented to participate in the experiment and those whose parents did not consent.

[17]For budgeting reasons we did not have a pure control group in which no student was treated. We discuss the implication of this in section 5.1.

[18]It is possible that teachers could have inferred which students were in the treatment group. We think this is unlikely given the many responsibilities teachers have for classroom activities and the size of classes in our school settings.

before the winter break and lasted through December 2014, picking up again in March 2015 and lasting until December 2015. The summer break happened from mid-December 2014 to early March 2015.

Figure 2: Timeline



Note: The figure shows the timeline of the intervention and data collection implemented in 2014 and 2015.

All parents in the treatment group received weekly messages on attendance, and monthly messages on classroom behavior and math test scores (separately).[19] We told parents how many days the child had attended school out of the previous school week (usually five days), and we provided parents with the number of positive, neutral, and negative classroom behaviors that teachers had recorded in the classroom notebook over the prior month. We provided monthly updates on the record of all math test scores in the semester, the average of these scores, and the classroom average score for the same tests. Hence, parents learned information about their own child's math performance, as well as how their child performed relative to the classroom average. In addition, parents in both the treatment and the control group received text messages about school meetings, holidays, and other general school matters throughout the year. We refer to these as "general" messages.[20] Parents of students in the control group continued learning about their child's academic performance through report cards that were sent home every quarter.

To create the information for these messages, we collected data on attendance, grades, and behavior from school classroom books. Our research team scanned and entered these data into a digital platform, which then automated the sending of messages each week. We sent more than $44,000$ text messages over 18 months: 68% provided information on attendance, 16% on math grades, and 16% on classroom behavior.[21]

[19]This differs from Bergman & Chan (2021), who only send text messages to alert parents of missing homework, tests, or classes.

[20]Online Appendix C explains in detail the intervention: production of messages, timeline, and delivery. It also provides a script of each type of message sent to parents.

[21]Behavior data were difficult to collect. In Chile, each classroom has a notebook in which teachers can make comments about particularly good or bad behaviors of specific students. For example, the teacher might

Our original research design included a complementary intervention to the text messaging treatment. The complementary investment consisted of a 9-minute parenting video that provided parents with advice on how to use the text message information provided by schools. In a random 50% of all classrooms, we allocated the parenting video to the text messaging treated parents only. This second treatment therefore worked as an add-on to the original text messaging treatment. We discuss the implications of this add-on treatment for empirical strategy and interpretation of results in section 5.1 and Online Appendix D. Online Appendix D details the various implementation challenges in the field that led to very few parents watching the video. In the rest of this paper, we focus on estimating the effects of the text messaging treatment (with or without the add-on parenting video).

4 Data

Data Sources. We use information from four data sources. First, we collected data on all students' math grades, daily attendance, and all behavior notes from classroom books for the years 2014 and 2015. These are daily-, weekly- and monthly-frequency data that we aggregate to an annual level. Second, we use student-level records provided by the central Ministry of Education of Chile. These records contain information on students' end-of-year school performance, including test scores, annual attendance rate, and grade retention, as well as basic demographic information. They are available for our sample of schools for the period from 2013 to 2015 and are used for allocating funding/subsidies across school. We use the 2013 data as pre-treatment controls and to generate our measure of students who are *at-risk* at baseline, and 2014 and 2015 Ministry data to validate our main results using classroom records. Third, we recorded all text messages' information such as day and time stamps, the messages' content, the name of the recipient parent, and the delivery status of the text message (i.e., whether the phone number received the message). Fourth, we administered several surveys to all parents and children participating in the experiment. Surveys were administered before the intervention took place (baseline), at the end of the first academic year (midline), and at the end of the second academic year (endline). Student surveys were conducted in class while parent surveys were sent home with children, who were encouraged to ask their parents to complete and return the surveys.[22]

Outcome Variables.[23] We use data recorded by teachers in classroom books to mea-

write, "Samuel concentrated well in reading," or "Taryn hit her friend during math class." We developed a system for categorizing such behavior "notes" as positive or negative, and followed these definitions in all classrooms.

[22]Online Appendix E provides more details and information on these data sources.

[23]Online Appendix F describes in detail each of the outcome and control variables used in this paper. It

sure our primary student outcomes: math grades, attendance rates and classroom behavior, which we aggregate at the annual level. Using administrative school records, we also measure outcome variables (i.e., grades, attendance rates, and an indicator for whether the student passed the grade) at an annual frequency at the end of each school year to validate our main sources.[24] Using classroom books we also constructed monthly math grades, attendance rates, and behavioral notes. All math grades were standardized using the corresponding grade-year control mean and standard deviation.[25] In addition, we built two indicator variables for meaningful thresholds required to pass the grade: 85% of annual attendance for passing the grade, and the 4.0 math grade for passing the subject. Using classroom books, we also measured negative behavior by adding all the behavioral entries during the school year (post-treatment) and then standardized the sum using the grade-year control distribution.

Our secondary outcome variables were designed to capture information gaps and certain behavioral responses to the treatment among students and parents. First, we built measures of information gaps by comparing survey questions that asked parents about their children's recent grades, absences, and behavior. We then compared parents' responses to students' responses and administrative records. These measures help us to test whether the text messaging treatment improved parent-school communication at all. Second, we asked parents and children a series of questions to compute pre-specified measures (i.e., several items that are aggregated into one variable usually referred to as a "scale") of study habits, academic efficiency, parental support, parental supervision, parental school involvement, and parental positive reinforcement. These were intended to capture any changes in home behaviors and parent-child or parent-school relationships that might result from the intervention. We administered a set of survey items from three sources: the University of Chicago Consortium on Chicago School Research; the Manual for the Patterns of Adaptive Learning Scales (PALS) developed by the University of Michigan; and scales on positive parenting developed by the Prevention Group at Arizona State University. We aggregated categorical answers into scales using a maximum likelihood principal components estimator. We then standardized answers using the mean and standard deviation of the control group. Overall, we find that each scale has good psychometric properties.[26] We asked parents and their children a similar set of questions. Scales are highly correlated both across survey waves and between children and parents –further suggesting that the quality of these scales is high (See Tables F.5 and F.6).

shows the specific data sources and provides a description of how the variables were constructed.

[24]We relegate most of the results using these data to the appendix tables and figures.

[25]In computing the control mean and standard deviations we only use information of the students that consented to participate in the study.

[26]Online Appendix F.1 describes how the scales were built. For both parents and students, we show the eigenvalue of each latent factor, the loading associated with each variable, and the Cronbach's alpha for each survey wave.

Finally, to assess how much parents value the information provided through our intervention, follow-up surveys asked parents about their willingness to pay for the text messages.[27] Parents were randomly assigned a value $V of (low) $500 Chilean pesos, (medium) $1000 and (high) $1500 price (where $ is Chilean pesos per month, and where $1,000 is about USD 1.50).

At-risk index. We build an index to measure each student's risk of failing classes or dropping out later in life. Specifically, we rely on three variables measured before the intervention began: standardized attendance ($Z_i^{attendance}$), math grades (Z_i^{grades}), and negative behavioral notes ($Z_i^{behavior}$).[28] The at-risk index is then defined as a simple average of these measures ($at-risk\ index_i = (-Z_i^{attendance} - Z_i^{grades} + Z_i^{behavior})/3$) which we standardize to the control group. The higher the value of this index, the worse grades, worse attendance, and worse classroom behavior the student has at baseline. Throughout the analysis, we rely on this index to assess the differential impact of the intervention on the primary and secondary outcomes for students with different values of the index.[29]

In our setting, low attendance and low grades are early warning signals for future grade retention and dropout. To explore this empirically, we used data from the Ministry of Education to look at the complete educational trajectory of almost 1.3 million students who were in grades 8-12 in the period 2006-2013 attending schools in the metropolitan area of Santiago. We estimated a simple model in which the dependent variable was an indicator for having being retained in the same grade or having dropped out of school and the independent variables were the attendance and GPA in the previous three years (two of three components of the at-risk index that we observe for the whole population). We find that all coefficients are negative and most are statistically significant at normal levels.[30]

Response rates. Baseline data from administrative sources are available for all students in the experimental sample (except for a handful of students who joined the schools mid-year in 2014). Administrative data are also complete for the first year of the experiment. During the second year of the experiment, due to the normal churn of students changing schools, we have information for 90% of the students. This attrition rate is similar for treated and control

[27]We asked: *"It is possible that next year your daughter's/son's school can send you regular text messages with information about their school performance (attendance, grades, and classroom behavior) four times a month. However, there might not be enough funds to provide this service free of charge. Thinking about how valuable this service would be for you, please tell us whether you will be willing to pay $V pesos a month to receive four text messages a month, from April to December."*

[28]We use final attendance and math grades from the academic year prior to the beginning of the intervention and accumulated negative behavioral marks during the month prior to the start of the intervention.

[29]From the onset of the experiment we set out to study differential treatment effects for students of different baseline achievement (attendance, grades, behavior). We did not, however, pre-specify the at-risk index or the heterogeneity analysis directly based on it.

[30]See Appendix Table 1.

students. Regarding survey data, students' response rates were between 91%, 89% and 80% across baseline, midline, and endline. More data were missing for parents, particularly from follow-up surveys. Parental response rates were 73%, 57%, and 54% at baseline, midline, and endline. For all survey waves, response rates were similar across treated and control students and parents. In addition, respondents may have chosen to complete some items but not others. This item non-response affects the sample sizes of secondary outcomes measured through the midline and endline parents' and students' surveys.[31]

5 Estimation and Experimental Validity

5.1 Empirical strategy

Intention-to-treat effects (ITT). To identify the effect of sending parents high-frequency academic information on students' and parents' outcomes we pool the two school years of the intervention and estimate individual-level regressions of the form:

$$Y_{ict} = \alpha_1 + \beta_1 T_{ic} + \psi_1 X^0_{ic} + \gamma_c + \pi_t + \epsilon_{ict} \tag{1}$$

where Y_{ict} is the outcome of student (or parent) i in classroom c of school j, and year t; T_{ic} is an indicator for whether a child's parents were part of the randomized group that received the information treatment, and it is constant over time; and π_t are year fixed effects. X^0_{ic} are the baseline standardized math grade and attendance rate.[32] Finally, γ_c are classroom-level fixed effects (strata in the experimental design). Despite the main randomized variation being at the student level, to be conservative, we cluster standard errors at the classroom level.[33] β_1 captures the intention-to-treat effect of the information sent by text messages. Because we include classroom-level fixed effects (γ_c), β_1 is identified through differences in individual-level treatment status within each classroom.[34]

[31] Online Appendix G shows the response rates for the different samples, years, and data sources. It also describes attrition from and entry into the sample, and the characteristics of those students in terms of their treatment status.

[32] For a handful of students baseline values are missing. In those cases, we impute the control baseline variables using the classroom-level mean. We add an indicator variable in the regression model equal to one for these observations.

[33] A classroom c is a unique combination of school, grade-level, and classroom in the first year of the intervention.

[34] As mentioned in Section 3, our original research design had a complementary investment (a parenting video) randomized to half of the classrooms. Within the video-treated classrooms, only parents that were already receiving text messages received the video. This implies that the parameter β_1 in equation (1) can in principle be capturing two effects: the treatment effect of the text messages and the treatment effect of the parenting video intervention times the probability of receiving that parenting intervention. In the Online Appendix D we discuss the parenting intervention, the research design, challenges with implementation

Classroom-level spillover effects. We exploit the differential classroom-level exposure to treatment to estimate spillover effects of the intervention on the treated. Such spillovers could be important, especially if such parent-school communication programs scale up to cover all enrolled students (rather than just a randomly selected treatment group), where by definition there would be no control group. Let E_c be an indicator variable equal to one if classroom c was randomized to have 75% of students treated and is equal to zero if it was randomized to have 25% of students treated instead. We estimate the parameters of the following model:

$$Y_{ict} = \alpha_2 + \beta_2 T_{ic} + \eta_2 T_{ic} \times E_c + \psi_2 X_{ic}^0 + \lambda_c + \omega_t + \varepsilon_{ict} \quad (2)$$

The coefficient η_2 measures the differential treatment effect of the text-message intervention in classrooms where a larger proportion of students was treated. Because of randomization, and assuming there are either no spillovers to the control group, or equal spillovers to the control group in all classrooms, η_2's estimate allows us to quantify the size of the spillover effect *on the treated students*. This is the relevant group when thinking about scaling the program to cover all students. In our experimental design E_c is collinear with λ_c, so we cannot estimate differential spillovers among students who were randomized out of the text messages treatment.[35]

If there are any positive spillover effects to the control group, such as those found by Bettinger et al. (2021), our treatment effect estimates ($\hat{\beta}_2$) would capture a lower bound of the effect of text messages on all students' outcomes. Moreover, as long as any spillovers on the non-treated are larger in classrooms where a higher share of students were treated, then our estimated spillover effects on the treated ($\hat{\eta_2}$) would also represent a lower bound of the true spillover effect to this group.[36]

which meant very few parents watched the video, the results, and the implications for the interpretation of the parameters in equation (1). We show evidence that our estimated $\hat{\beta}_1$ is mostly capturing the treatment effects of the text message intervention.

[35]Estimating model (2) without classroom fixed effects would not respect the research design, and would not allow us to control for variations in class size (in our sample, classes vary from 20 to 44), consent rates across classrooms (mean consent rate is 54%), and possibly other classroom characteristics not observable in the data. This could affect the estimated treatment effect if the number of treated students has an additive impact.

[36]The assumption that there is a dose-response relationship between the size of the share of students treated in the same classroom and the spillover to the control group is a reasonable one. Avvisati et al. (2014) provide evidence consistent with spillovers increasing with the level of interaction between treated and non-treated students in the same classrooms in their parent-school intervention in French middle schools.

5.2 Balance on Pre-Treatment Observable Characteristics

We compare the observable characteristics of students and parents assigned to the treatment and control groups before the intervention began.

Table 1 shows total observations with available data (column 1)[37]; the average of each variable for the treatment group (column 2) and the control group (column 3); and the p-value of the null hypothesis that, conditioning on classroom (strata) fixed effects, the differences between treatment and control averages are zero (column 4).[38]

Panel A shows statistics based on administrative records. In our sample, 45% of students are female. The median age is 9.8 years. Students in treatment and control groups have similar grades at baseline, with math and language scores around 5.1 (on a 1-7 scale), similar attendance rates (89 percent), and similar levels of the at-risk index. About 95% passed their grade in the year prior to the experiment. Pre-treatment administrative records are missing for about 9 percent of the sample. We cannot reject equality between any of the mean characteristics of students randomized to treatment and control. The last row of the panel presents a Wald test of the joint null hypotheses that the differences in means reported in columns 2 and 3 for all the variables in the panel are zero. We cannot reject that null at standard levels of significance.

Panels B and C show standardized parents' and students' scales from the baseline surveys.[39] Before the intervention began, students in the treatment and control groups reported putting in similar effort when studying at home, received the same parental supervision, involvement in their school affairs, and positive reinforcement at home. Parents across treatment and control groups similarly report the same parenting practices at home. We reject equality at the 10 percent level for one measure with parents in the treatment group reporting less family support than parents in the control group. Despite not rejecting most of the null hypotheses that the average scales are similar for treated and control students, we note that in most cases the estimated means are lower in the treatment group. This could reflect the fact that many of these scales could be noisy measures of a similar latent variable.

[37] We note that the number of observations vary throughout the manuscript for three reasons: there are two samples (main sample and full sample), which we analyzed in two formats (cross-sectional versus panel data analyses), and are sometimes affected by non-response (both survey non-response and item non-response).

[38] Panel A of Appendix Figure 1 shows that observable characteristics are similar between treatment and control students when the full sample is used or in the sample of respondents to the parent's and student's baseline surveys. Additionally, Panel B reports a similar balance table to that shown in Table 1; it includes an additional variable to indicate whether the classroom was randomized to receive a high or low share of treatment, and the interaction with T_{ic}.

[39] The survey items used to build these scales can be found in Online Appendix Tables F.2 and F.3.

Table 1: Students' and Parents' Pre-Treatment Characteristics

	Obs.	Treatment Mean (μ_T)	Control Mean (μ_C)	p-value of adj. dif.
	[1]	[2]	[3]	[4]
Panel A: Administrative records				
Female	1066	0.45	0.47	0.57
Age	1066	9.81	9.79	0.41
New student	1066	0.08	0.07	0.42
Language grade	976	5.10	5.07	0.85
Math grade	976	5.14	5.19	0.37
Final avg. grade	976	5.57	5.59	0.47
Attendance rate	976	0.89	0.89	0.53
Passed grade	1018	0.95	0.96	0.57
At-risk index (standardized)	1066	0.05	0.00	0.35
Missing grades/attendance/pass data	1066	0.09	0.08	0.41
Multiple hypotheses Wald test				0.72
Panel B: Parents' Survey Data				
Standardized scales ($\mu_c = 0$, $\sigma_C = 1$)				
Study habits	704	–0.07	0.00	0.51
Academic efficiency	730	–0.09	0.00	0.16
Family Support	739	–0.12	0.00	0.06
Low Family Supervision	709	–0.06	0.00	0.72
Parent School Involvement	716	–0.01	0.00	0.66
Positive reinforcement	738	–0.06	0.00	0.31
Parent scales index	773	–0.06	0.00	0.21
Mother completed high school	774	0.53	0.49	0.78
Missing baseline survey	1066	0.26	0.27	0.59
Multiple hypotheses Wald test				0.39
Panel C: Students' Survey Data				
Standardized scales ($\mu_c = 0$, $\sigma_C = 1$)				
Study habits	909	–0.19	0.00	0.10
Academic efficiency	915	–0.14	0.00	0.15
Family Support	864	–0.15	0.00	0.12
Low Family Supervision	859	0.05	0.00	0.60
Parent School Involvement	858	–0.12	0.00	0.59
Positive reinforcement	868	–0.04	0.00	0.90
Student scales index	962	–0.17	0.00	0.15
Missing baseline survey	1066	0.08	0.09	0.84
Multiple hypotheses Wald test				0.11

Note: Column [1] shows the number of observations with non-missing data, columns [2] and [3] the mean value of each baseline characteristic in the treated and control group, respectively. Column [4] reports the p-value on the treatment coefficient in a regression using each baseline characteristic as the dependent variable. All tests adjust for classroom fixed effects and robust standard errors are clustered at this level. Parent and student scales index are simple scales' averages which were standardized using the control mean and standard deviation so that standardized scales for the control group have a mean $\mu_C = 0$ and a standard deviation $\sigma_C = 1$. Observable variables in Panel A correspond to 2013 except for new student variable that refers to 2014. The rows "Multiple hypotheses Wald test" reports the p-value of a joint test of the null that all the differences in means of the variables reported in each panel (of treated and control students) are zero. We exclude from this test the variable that reports the proportion of missing observations.

For this reason, we aggregated all these scales into parents' and students' indices.[40] We cannot reject equality of the mean of the indices of treated and control students. Finally, we find that mothers in the treatment and control groups are equally likely to have completed high school. The final row in each panel presents the p-value of the joint test of equality of the variables listed in the panel. In both cases we cannot reject the null at the 10 percent confidence level.

5.3 Delivery of Text Messages

All text messages were sent to parents as planned. However, not all text messages were actually received.[41] Several factors contributed to reception failure. A message was more likely to fail if the network was very busy, if some technical problem surfaced within the network, or if a parent had changed their phone number during the experiment. To maximize the chances that text messages reached parents, we sent the messages on Mondays, when the network was not as busy as on other days.[42] At the beginning of the second school year during which the experiment took place, we also recontacted all consenting parents to verify or update their cellphone numbers.

Table 2 shows estimates obtained with equation (1) where the dependent variable is the total number of messages sent (first row) or received (second row) during the course of the experiment. The variables are computed for each type of message (attendance, grades, classroom behavior, general, and all) using information from the digital platform described in Section 3. Each point estimate shows the coefficient estimate of β_1, which estimates the differences in the total number of text messages sent to/received by parents in the treatment group and those in the control group.

By the end of 2015, when the experiment had run for one and a half school years, an average of 44 more text messages per year had been sent to parents in the treatment group than to parents in the control group. Over the same period, an average of 26 messages per year had been received by parents in the treatment group. This implies that almost 60% of sent text messages were successfully received by the end of the intervention, a success rate similar to those reported in the literature. Bergman & Chan (2021), for instance, report that in their text messaging information intervention in West Virginia, about one third of treated parents never received messages that were sent.

The bottom panel shows the distribution of messages sent and received for parents in the

[40] To compute the parents'/students' scales index we added all the standardized scales with a positive connotation and subtracted the low family supervision scale. We then normalized by the number of scales and standardized using the control group's mean and standard deviation.

[41] After sending a text message, cellphone companies mark that message as received or failed to be sent.

[42] During the first two months of the experiment, messages were sent on Fridays.

Table 2: Compliance by Type of Text Message

	All	Attendance	Behavior	Grades	General
	[1]	[2]	[3]	[4]	[5]
Panel A: Text messages sent					
T	43.960***	29.966***	6.715***	7.326***	–0.047
	[0.704]	[0.447]	[0.085]	[0.130]	[0.079]
Panel B: Text messages received					
T	26.341***	17.646***	4.506***	4.337***	–0.148
	[0.777]	[0.452]	[0.122]	[0.127]	[0.123]
Observations	2011	2011	2011	2011	2011
Control mean messages sent	5.520	0.000	0.000	0.000	5.520
Control mean messages received	3.741	0.000	0.000	0.000	3.741
% text messages received / sent (among treated)	0.645	0.623	0.634	0.632	0.638
Proportion of messages across type (sent)		0.549	0.123	0.131	0.198
Proportion of messages across type (received)		0.527	0.133	0.126	0.213

Note: "Text messages sent" refers to the cumulative number of text messages sent to student's parents. "Text messages received" refers to the cumulative number of text messages with a confirmed delivery status. Columns [2]-[5] report the T_{ic} coefficient of equation (1) with the annual number of each type of text message as the dependent variable. Column [1] adds all types of text messages. Attendance, grades, and classroom behavior text messages were sent only to the treatment group. General text messages were sent to all treatment and control individuals. All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

treatment and control groups. Most of the messages were about attendance, because these were sent weekly, while classroom behavior and grade messages were sent monthly. These treatment messages were only sent to, and received by, parents assigned to the treatment group. By contrast, parents of students in the control group were sent (and received) general text messages at largely the same rate as those in the treatment group (column 5).[43]

The data suggest that the probability of receiving text messages is unlikely to be correlated with family-level characteristics that also affect child outcomes of interest. We might worry, for instance, that parents who have low attachment to the labor market and unstable incomes are also more likely to switch cell numbers. They would then be less likely to receive text messages about their children's academic performance. Children in these families may also have worse school outcomes. To assess this possibility, we estimated a regression model in which the dependent variable was the total share of successfully delivered text messages (total received/total sent) on baseline attendance and math grades, age, gender, a composite

[43] Panels A and B of Appendix Table 2 reports the treatment compliance in each year of the intervention (2014 and 2015). More messages were sent in 2015, when the intervention was implemented for a full school year, than in 2014 when the intervention was implemented during the second half of the school year. Panel C presents the compliance for the full sample.

index of the parent scales and mother's education (as reported in Table 1), and classroom fixed effects. Students with higher baseline grades, with higher attendance, or with higher family support and supervision are no more (or less) likely to receive text messages. Mother's education seems to be weakly correlated with the share of messages received.[44].

Beyond the matter of whether parents received text messages that were sent, there is also the question of whether parents read the text of the messages that they received. In the follow-up surveys we asked parents if they had received text messages with information on their children's school outcomes. We found that parents in the treatment group were more likely to answer that they had received text messages regarding their child's attendance, grades, and classroom behavior.[45]

6 Results

6.1 Main Results: Students' Academic Outcomes Improved

Table 3 presents the main results of our paper. We show the estimates of the intention-to-treat effects (using equation 1) of the intervention on our primary students' outcomes measured using classroom books: standardized math-grade outcomes at the end of each year (column 1), an indicator for whether the annual math grade was a passing grade (above 4.0) (column 2), yearly attendance rate (column 3) for each year, an indicator for whether attendance was above the 85% cutoff required for the student to pass the grade (column 4), and standardized total annual negative behavioral notes (column 5).

The ITT estimates show positive and significant effects on students' school performance. Math grades improved by 0.088 of a standard deviation. This positive impact on math grades pushed more students over the 4.0 cutoff for passing the subject, increasing this probability by 2.7 percentage points. The treatment also improved attendance by almost 1.1 percentage point leading to a 4.7 percentage point increase in the number of students who met the 85% attendance rate threshold needed to pass the grade.[46] On average, the treatment did not have an impact on the occurrence of negative classroom behaviors.

Our main results are robust across a range of different specifications, sample choices, and data sources. Appendix Figure 2 presents results from estimating the effects of the

[44] In Online Appendix C.3 we present and discuss these results. We also show that people who received the text messages (compliers) are very similar to those that were sent text messages but did not received them (non-compliers) based on a wide set of pre-treatment variables.

[45] See Panel D of Appendix Table 2.

[46] Larger treatment effects in column 4 compared to column 3 suggests the possibility of bunching around the threshold. We tested for a discontinuity in the attendance distribution in the year prior to the intervention following Cattaneo et al. (2018). We rejected the null that the distribution is continuous (p-value=0.03).

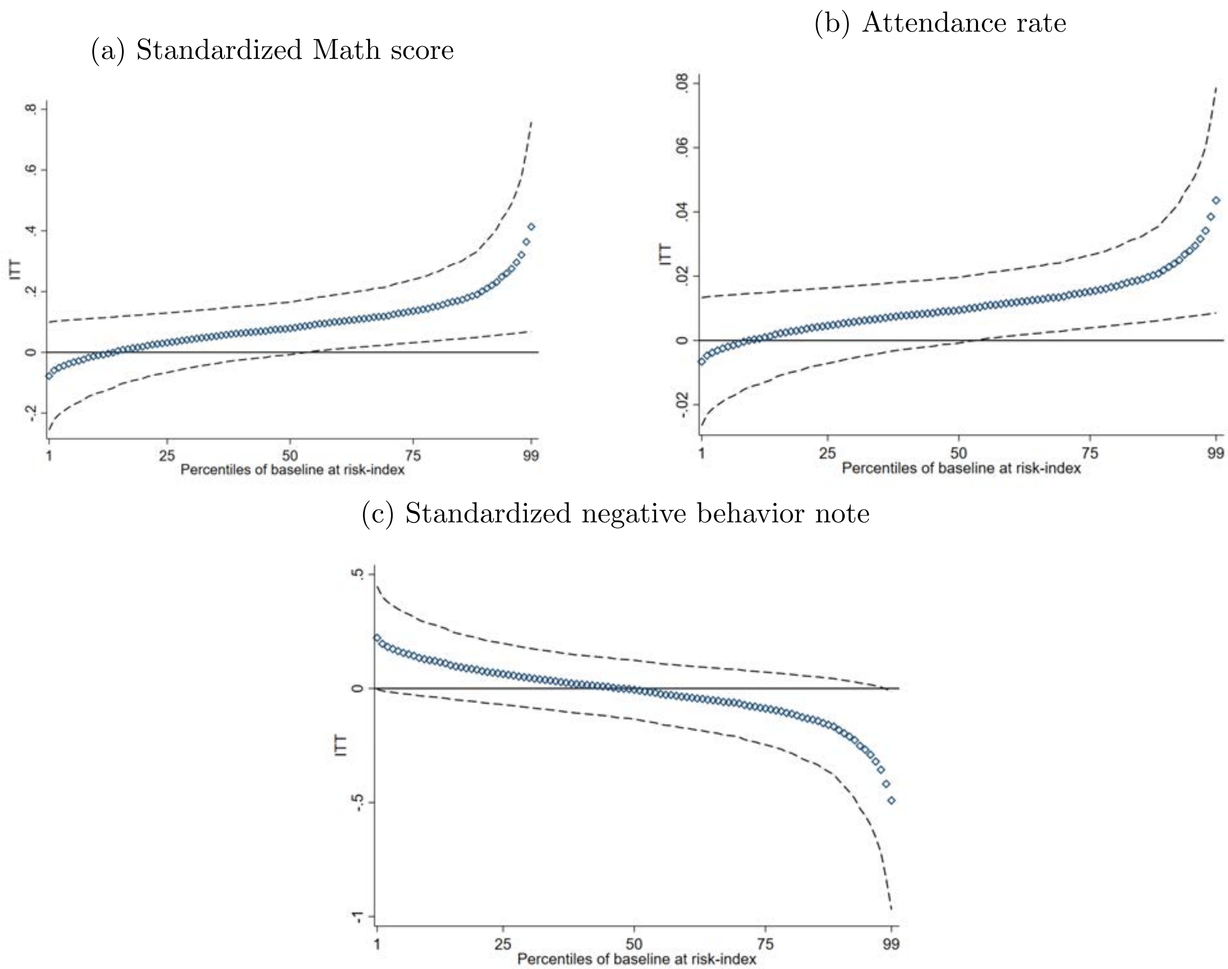
Table 3: Treatment Effects on Grades, Attendance and Classroom Behavior

	Standardized math grade	Math grade >4.0	Attendance rate	Cumulative attendance >85%	Standardized # negative beh. notes
	[1]	[2]	[3]	[4]	[5]
Panel A: Treatment Effects					
T	0.088*	0.027**	0.011**	0.047*	0.004
	[0.045]	[0.013]	[0.005]	[0.024]	[0.075]
Panel B: Heterogeneity					
T	0.088*	0.026*	0.010*	0.047*	–0.019
	[0.044]	[0.013]	[0.005]	[0.024]	[0.067]
T x at-risk index	0.140*	0.025	0.014*	0.073**	–0.203**
	[0.071]	[0.019]	[0.007]	[0.028]	[0.094]
Observations	2011	2011	2011	2011	2011
Control mean	0.00	0.934	0.877	0.728	0.00

Note: Panel A shows the intention-to-treat (T) estimates and its corresponding standard error estimated using equation (1) using OLS. Panel B adds the interaction with the student-level at risk index. At-risk index is a simple average of standardized baseline attendance, math grades and negative behavioral notes. All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Models in Panel B additionally include the at-risk index variable as control. Columns 1 and 5 report results on outcomes that were standardized so that mean among the control students is zero and the standard deviation is one. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

treatment on grades, attendance, and behavior for specifications that include and exclude baseline controls; that separate out the midline and endline samples; for samples that include students who leave the study in year two (either because they are in grade 8 in the first year or attend the one school that dropped out of our study at the end of year one); and that use outcomes data from the national ministry rather than the administrative data collected by our research team directly from schools. While the effects on math grades are larger in 2014, the impact on attendance rates appears to be stronger in the the second year of the intervention. Overall, while the confidence intervals move around somewhat with different choices of samples and outcomes, the point estimates for the impacts of the treatment on grades and attendance are uniformly positive. The main results in our Table 3 are in the middle of the range of estimates in Appendix Figure 2. And, for each outcome, we could not reject the hypothesis that the point estimates are the same across different samples, specifications, and source of outcomes data, and the same as in Table 3. The fact that the treatment produces stable positive impacts on our main grade and attendance outcomes is reassuring.[47]

Figure 3: Predicted Treatment Effect by baseline at-risk index



Note: Figure shows linear predictions and 95% confidence intervals of the intention-to-treat (ITT) estimates on math grades, attendance rate and negative behavior. Computed based on coefficients from columns [1], [3] and [5] of Table 3 panel B, respectively. The standard error for estimate at each percentile p is constructed as $\sqrt{Var(\hat{\delta} + \hat{\beta}_Z \times \overline{Z}_p)}$, where $\overline{Z}_p$ is the mean of at-risk index in percentile p.

Panel B of Table 3 shows estimates for students with different pre-treatment risk of failing grades or poor attendance. To estimate these effects, we interacted the at-risk index described in Section 4 with the randomized treatment indicator variable (in equation 1) and

[47] We account for the imperfect compliance with treatment by estimating local average treatment effects. Let D_{ic} be an indicator variable equal to one for those treated students whose parents received at least one text message with information on each specific outcome (i.e., compliers). We then include D_{ic} –instead of T_{ic}– in equation (1) which we instrument in a first stage with the randomized treatment variable T_{ic}. Appendix Table 3 shows the results. Point estimates are larger in absolute value than those presented in Table 3; they are values scaled-up by the proportion of parents who actually received the text messages. These results are robust to other definitions of compliance with treatment like, for instance, having received more than 75 percent of the messages.

controlled for the at-risk index. The intervention had the largest impacts on math grades, attendance and improvements in behavior for students who were more at risk before the intervention started. The treatment effects are two to three times larger for students with an at-risk index one standard deviation larger than the mean (which by construction of the index is zero for the control group). Figure 3 explores this result in more detail by plotting the linear prediction of the treatment effects on math grades (Panel A), attendance rates (Panel B), and classroom behavior (Panel C) for students with different levels of the at-risk index. We find that effects for attendance and math grades are larger and statically significant only for students at higher risk. The pattern of behavioral effects by the at-risk index also suggest larger improvements (less negative behavior notes) for students most at-risk, although the confidence intervals in Figure 3 Panel C cannot reject zero. Note that the results in Table 3 Panel B are consistent with the treatment increasing the probability of the most at-risk students achieving the attendance and math grades thresholds for passing the grade and subject; precisely for the population of students who have a higher probability of dropping out in later years. Improving parent-school communication through this text messaging program seems to naturally target, and improve outcomes for, students who need the most support in school and at home.

6.2 Classroom-Level Spillovers on the Treated

In the presence of treatment spillovers among the treated, the treatment effect could vary with the share of other treated students in the classroom. This could happen, for example, if the value of skipping school falls when friends are no longer truant (Bennett & Bergman 2021). Alternatively, if a student's friends are working harder to improve their grades, that student's own effort to earn better grades may increase (if, for instance, there are ranking concerns (Tincani 2018)). Spillover effects are important to quantify when considering possible impacts at scale. To estimate these indirect effects of the intervention, under the assumptions discussed in Section 5.1, we exploit the randomization of the different shares of students who were part of the treatment group in each classroom.

Table 4 presents the results of the ITT spillovers for the same set of outcomes as in Table 3. Note that the interaction coefficient captures the *differential* effect of the spillovers by comparing classrooms with high and low shares of treated students; in other words, it examines whether there is extra value evident in being in the text messaging program when many more classmates are also in the program. In all cases, although point estimates are imprecise and not statistically significant in columns 1 and 3, the *differential* effect of being assigned to treatment in a high-share classroom *improves* educational outcomes of treated

Table 4: Spillover Effects

	Standardized math grade	Math grade >4.0	Attendance rate	Cumulative attendance >85%	Standardized # negative beh. notes
	[1]	[2]	[3]	[4]	[5]
T	0.070	0.006	0.005	0.009	0.113
	[0.054]	[0.015]	[0.007]	[0.034]	[0.095]
T x High-Share	0.042	0.052*	0.013	0.091*	–0.258*
	[0.094]	[0.027]	[0.011]	[0.046]	[0.150]
Observations	2011	2011	2011	2011	2011
Control mean	0.00	0.934	0.877	0.728	0.00
p-value $H_0: T + T \times H = 0$	0.15	0.01	0.02	0.00	0.22

Note: Each row shows the intention-to-treat estimates and its corresponding standard error estimated using equation (2) using OLS. T refers to the randomized individual-level treatment (equal to 1 if parents were sent text messages and zero otherwise). $High - Share$ refers to the randomized classroom-level treatment (equal to 1 for high-share classrooms and zero for low-share classrooms). All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Columns 1 and 5 report results on outcomes that were standardized so that mean among the control students is zero and the standard deviation is one. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

students –it is larger than the main effect of the treatment in low-share classrooms. The last row presents the p-value of the null hypothesis that the treatment effect was zero in high-share classrooms which is rejected at the 10% level in columns 2, 3 and 4. This suggests positive spillovers of the intervention among treated students. With a higher share of treated peers, students are significantly more likely to meet the 4.0 passing grade cutoff and to reach the 85% attendance cutoff.

The spillover results in Table 4 suggest that we would not expect any negative impacts of scaling up this intervention to cover all students. If anything, we should expect even larger impacts at scale, when everyone is treated.

6.3 Do Text Messages Work in the Same Way Over Time?

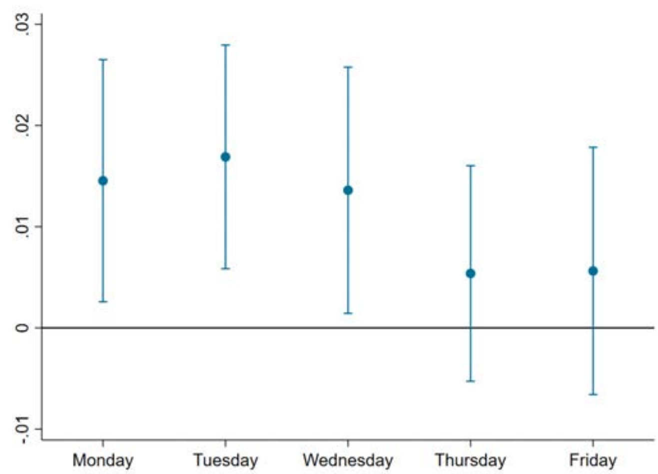
Bergman & Chan (2021) note that there are many open questions about how parents will respond to ongoing text-messaging from schools. The long duration of our treatment intervention allows us to explore how parents responded to the text messaging over time.

Parents who receive text messages might forget about the content of the messages after some time, and this could affect their decisions about whether to allow their children to miss a day of school. The majority of the weekly attendance text messages were sent on Mondays.

We use daily attendance data to explore whether the effectiveness of the text messages fades within the week.[48]

Figure 4 depicts point estimates and confidence intervals for models similar to that of equation (1), which was modified to include an interaction of the share of text messages received with days-of-the-week indicator variables. We find a pattern suggestive of fade out over the week. Attendance by students in the treated group is significantly higher than attendance of students in the control group on Mondays, Tuesdays and Wednesdays; by contrast, attendance rates of the two groups are indistinguishable on Thursdays and Fridays.[49] However we cannot reject equality of the coefficient estimates. Rogers & Feller (2018) find similar results with a larger impact in the week immediately following the delivery of the treatment. This result suggests that the treatment effect of the text messages could be somewhat short-lived. Information treatments delivering information that depreciates in value over time may need to be high frequency in order to be effective.

Figure 4: Weekly Fade-out of Attendance Treatment Effects



Note: Coefficients are obtained from the daily intention-to-treat estimates of Appendix Table 4. Standard errors clustered at the classroom level. Confidence intervals are at the 90% level.

A related concern is that parents could at some point stop paying attention to the content of the communication, or stop internalizing the information after having received such messages over some period of time. Because our intervention lasted for one and a half school

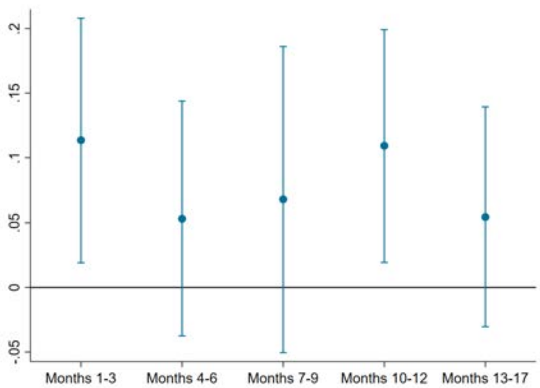
[48] After the first two months of the intervention, we started to systematically send all the text messages on Mondays. For this part of the analysis, we restrict the sample to this period and keep only observations for those students whose parents were sent and actually received the messages on Monday.

[49] Appendix Table 4 shows the estimated coefficients used to construct this figure and p-values of tests of equal coefficients. We reject the null that all coefficients in Figure 4 are equal (p-value=0.037) and the null that the treatment effect on Monday's attendance is equal to that of Friday's attendance -against the alternative that is lower- (p-value=0.065).

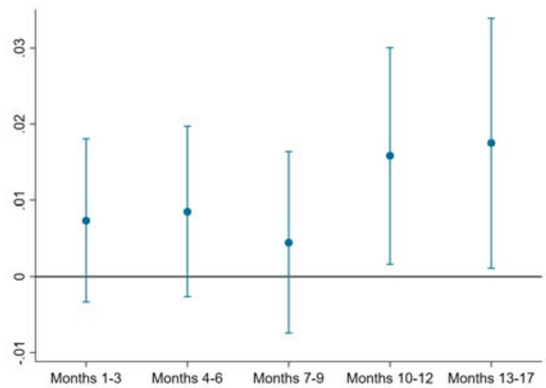
years we can explore the treatment effects over the months of the intervention. We estimated effects by month-groups interacting the treatment with month-groups identifying groups of months since the beginning of the intervention. Figure 5 plot the estimates and confidence intervals on the impact on monthly attendance, monthly math grades, and monthly negative behavioral notes.

Figure 5: Treatment Effects Over Time

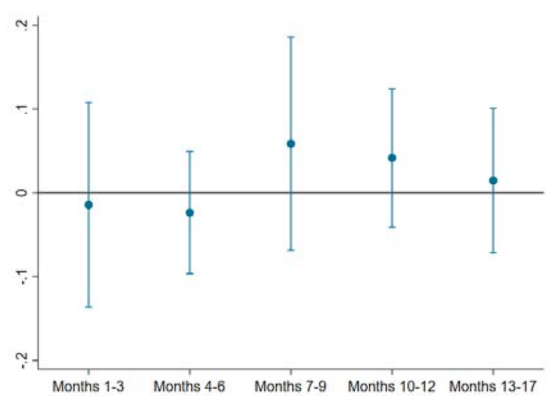
(a) Standardized monthly math grade



(b) Monthly attendance



(c) Standardized monthly negative behavior notes



Note: Coefficients are obtained from the respective intention-to-treat estimates of Appendix Table 5. Standard errors are clustered at the classroom level. Confidence intervals are at the 90% level.

We find that the impact on attendance is mainly concentrated in the last months of the intervention, although we cannot reject the null that all coefficients are equal.[50] In the case of math grades and behavior, there is no clear pattern in the timing of the effect. This is consistent with students/parents dynamically optimizing attendance behavior. The intervention could have more of an impact on absenteeism than grades by the end of the year

[50] Appendix Table 5 reports the estimated coefficients used to construct this figure. The p-values associated to the null of equality of the estimated coefficients are 0.766 (panel A), 0.751 (panel B), and 0.555 (panel C)

because that was when parents/students started to realize that the absences had accumulated enough to matter. It could also be the case that approaching the end of the school year, attendance is easier to move than test scores. From a policy perspective these results suggest that parents do not become immune to the intervention over the course of 18 months.

7 Did the Text Messages Intervention Improve Parent-School Info Gaps and Change Parenting Behaviors?

Our intervention was designed to close information gaps between parents and schools and promote parent engagement with students, and with schools. In this section, we explore some of these underlying mechanisms that might have contributed to why students' school performance improved after their parents were exposed to high-frequency text messages containing student-specific information. We show that the treatment was able to close existing parent-school information gaps about math grades, attendance and behavior while also improving parent attentiveness to other non-targeted aspects of school performance. The new information seemed to have changed the way parents provide support and supervise their children at home. All of these changes reflect greater parent engagement with day-to-day school activities of their children.

7.1 Parent-School Information Gaps Narrowed

We study whether the text messages reduced the prevailing parent information gaps regarding students' academic performance; to do this, we compare the accuracy of information among parents in the treated and control groups. We construct different measures of the accuracy of parent's beliefs regarding their child's school performance. Specifically, we contrast parents' responses with student surveys, classroom books, and school records. We then estimate treatment effects using equation (1), in which the outcome variables are the misinformation measures.

Table 5 presents the ITT effects.[51,52] Columns 1-2 measure parental misinformation regarding a student's attendance. Surveys asked parents about their child's absences with and without permission in the previous two weeks. We contrast parents' responses to students'

[51]The share of parents who are misinformed is larger for misbehavior than it is for attendance than it is for grades. This could be because the misbehavior and attendance are not reported in the students' report cards while grades are. In addition, each of the variables has a different range allowing parents more or less scope to make mistakes in their assessments.

[52]We computed the magnitude of the information gap for those parents without missing data. The average gap in attendance/grades is equivalent to 1/2 of a standard deviation in the attendance/grades distribution.

own responses on total absences (column 1) and to actual absences recorded in classroom books (column 2). Columns 3-4 assess the effect of the intervention on parental information about students' grades. Columns 5-6 capture parental misinformation about students' misbehavior. In both cases we also contrast parents' responses with students' surveys responses (column 3 and 5) and with classroom books (column 4 and 6). In all cases, the outcome variable is an indicator variable that is equal to one if the parent response does not match the student's responses or the administrative records.[53]

Table 5: Treatment Effects on Parental Misinformation

	Attendance Misinformation		Grades Misinformation		Behavior Misinformation	
	All absenteeism (Surveys)	All absenteeism (Admin.)	All grades (Surveys)	All grades (Admin.)	Misbehavior (Surveys)	Misbehavior (Admin.)
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Treatment Effects						
T	–0.079**	–0.014	–0.012	–0.027	–0.080**	–0.083**
	[0.039]	[0.039]	[0.045]	[0.036]	[0.034]	[0.038]
Panel B: Heterogeneity						
T	–0.082**	–0.011	–0.019	–0.029	–0.072**	–0.086**
	[0.040]	[0.039]	[0.048]	[0.037]	[0.033]	[0.038]
T x at-risk index	–0.012	0.035	–0.091	–0.021	0.081	–0.052
	[0.066]	[0.047]	[0.061]	[0.046]	[0.056]	[0.056]
Observations†	992	1143	827	1185	1140	1188
Control mean	0.535	0.392	0.398	0.319	0.639	0.470

Note: Panel A shows intention-to-treat (T) estimates and its corresponding standard error estimated using equation (1) using OLS. Panel B adds the interaction with the student-level at risk index. At-risk index is a simple average of standardized baseline attendance, math grades and negative behavioral notes. All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Models in Panel B additionally include the at-risk index variable as control. Column outcomes are indicator variables constructed by contrasting responses in parent surveys with those of student surveys or administrative records (shown in parentheses). Column [1] measures parental misinformation on all absenteeism (with and without parent permission in the previous two weeks) contrasting the responses of parents with those from students. Parents are classified as misinformed if they do not answer at least one of the questions, or if at least one of the answers (in bracket days) provided by students and parents do not match. Column [2] measures misinformation on all absenteeism (with and without permission) contrasting parent responses with classroom books. The ends of original bracket days in absences with and without permission are added to construct new bracket days. Parents are classified as misinformed if they do not answer at least one of the questions, or if classroom books' records of absences over the previous two weeks do not fall in the range. Column [3] contrasts parent and student responses and parents are classified as misinformed if they do not answer, or if reported grades' brackets do not match. Column [4] measures parental misinformation regarding all grades by contrasting parent responses about the student's last end-of-year grades with school records. Parents are treated as misinformed if they do not answer, or if the absolute difference between reported and actual grades is greater than 0.5. Columns [5] and [6] measure misinformation about student misbehavior by contrasting parent answers with student answers, and with information from classroom books, respectively. Using a four-value scale, parents and students were asked about the degree of agreement with the student's misbehavior statements. For column [5], parents are classified as misinformed if they do not answer at least one of the questions, or if the average absolute difference between parent and student answers are larger than the median (0.8). For column [6] parents are treated as misinformed if they do not answer; if the parent's average answer is equal to or larger than the median (2), and student did not misbehave according to classroom books; or if the parent's average answer is less than the median answer and student misbehaved in class according to books. Standard errors are clustered at the classroom level (shown in brackets).† Number of observations vary by column because of survey and item non-response. * significant at 10%; ** significant at 5%; *** significant at 1%.

Panel A of Table 5 shows that all point estimates are negative. That is, text messages reduced information gaps about student attendance, grades and classroom behavior. Parents' reports got closer both to students' reports and to school administrative records. Because our

[53] When comparing with classroom books, we allowed for a "mistake" of 1 absence and 0.5 points in the case of grades.

sample of parents who responded to the follow-up survey is relatively small, these reductions in information gaps are not always precisely estimated; nevertheless, coefficients are large and negative for all outcomes.[54] The ITT estimates, for instance, show that text messages significantly reduced the probability that parents misreported the number of their child's absences; the likelihood of such misreporting fell by 7.9 percentage points, in comparison to the results from student surveys. When we compare parents' beliefs with classroom books, the results also show a decline in information gaps, but not to a degree that is statistically significant.

In addition, the information intervention seems to have improved the accuracy of parents' knowledge of their child's grades. Although not statically significant at conventional levels, coefficients are negative and stable across outcomes. We also find a significant improvement in the precision of parents' assessment of their child's misbehavior at school. Overall, these results suggest that treated parents had more accurate information about their child's grades, attendance and classroom behavior after the treatment.

Panel B of Table 5 tests whether treatment effects on information gaps vary for students with different baseline values of the at-risk index. The intervention seems to have improved the accuracy of parents' beliefs about their child's grades and behavior for students with a higher at-risk index (although results are not statistically significant).

We notice that the impact of the treatment on grades in Table 3 was larger (in percent terms) than the impact on attendance in that table; however, the parent misinformation gap shrinks more for attendance than for grades. But attendance, grade, and behavior are measured in different units, and the range of possible parent responses to questions about these outcomes also differs from the range of actual outcomes (for details, see the table notes in Table 5). For this reason, and because our smaller parent sample makes precise estimation challenging, we view the results in Table 5 as broadly consistent with the view that exposure to the treatment improved parent information sets. Closing the information gaps was one channel through which the text message intervention improved schooling outcomes.

7.2 Effects on other subjects, and parent misinformation about those subjects

In Table 6 we estimate effects of the treatment on other, non-targeted subjects using outcomes data reported by the schools to the national ministry. We see that language scores increased by a significant 0.11 of a standard deviation, and scores on natural science and

[54] We cannot reject equality of treatment effects on information gaps based on students' reports and those based on administrative records.

history also increased by 0.05-0.09 of a standard deviation (not significant). This positive impact of the treatment on non-math subjects could have occurred through the channel of increased attendance (i.e. a positive downstream impact of the treatment). However, it might have also increased parental attention to school in general, thus leading to improvement in non-targeted academic subjects.

Table 6: Treatment effects on Other Subjects' grades and misinformation

	Language	Natural science	History
	[1]	[2]	[3]
Panel A: Standardized grades			
T	0.113*	0.098*	0.054
	[0.059]	[0.057]	[0.044]
Observations	1946	1916	1916
Control mean	0.00	0.00	0.00
Panel B: Misinformation			
T	–0.079**	–0.048	–0.054
	[0.033]	[0.044]	[0.041]
Observations	1142	973	972
Control mean	0.499	0.534	0.493

Note: Panel A and Panel B show intention-to-treat (T) estimates on subjects not targeted by the intervention. Panel A shows the effect on grades and Panel B on parental misinformation regarding those grades. Point estimates and standard error were estimated using equation (1) using OLS. All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata) and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Columns 1-3 of Panel A report results on outcomes that were standardized so that mean among the control students is zero and the standard deviation is one. Columns [1]-[3] of Panel B measure parental misinformation each subject grade. Parents are treated as misinformed if they do not answer or if the answered grade bracket does not match to the actual grade from administrative data. Standard errors are clustered at the classroom level (shown in parentheses). * significant at 10%; ** significant at 5%; *** significant at 1%.

In Panel B of the same table, we show some suggestive evidence that the treatment may have reduced parent misinformation in general. We estimate the impact of the treatment on parental misinformation about other subjects not specifically targeted by the intervention. Across the board, parent misinformation relative to the administrative records shrinks; the coefficients for parent information gaps about languages, social studies, and history are all negative. Interestingly, parent information gaps in languages shrink to about the same extent as they shrink for math grades (Table 5 column (1)). The results in Table 6 are consistent with two plausible explanations. First, in addition to reducing information gaps on the specific topics on which parents received information, the text message intervention could have induced parents to pay more attention to how their children are doing in other (non-math) subjects. Alternatively, because grades are correlated across subjects, parents updated their beliefs about their child's performance on the targeted subject (math) and, jointly, on all other subjects.[55]

7.3 Parent Engagement at Home and School Improved

By providing parents with information over a sustained period of time, the intervention may have led students and parents to respond with changes in behaviors at home – which, in turn, might then have resulted in better outcomes at school. To examine this, in Table 7 we analyze the responses to survey questions that were put to both parents (Panel A) and students (Panel B) in an identical manner. Columns 1 and 2 measure students' academic responses in terms of two aggregate scales (study habits and academic efficiency). Columns 3-6 looks at parents' behavioral responses, in terms of several aggregate scales designed to capture family support, supervision, involvement with school matters, and positive reinforcement at home.

These aggregate scales are built from individual survey items. Looking at control group means on these outcomes provides a clearer picture of the status quo. About 66 percent of students in the control group considered themselves to be organized with school work and 80 percent thought they were capable of understanding difficult school content. Approximately 93 percent of parents reported having shown to their children that they are proud of them and to have congratulated their child regarding school achievements. 36 percent reported that their children went to school alone and 29 percent reported a communication with a child's teacher[56]

We do not find a clear pattern or statistically significant results for the information provided by parents in terms of how the treatment affected their self-reported behaviors. By contrast, however, treated students perceived that they received significantly more family support as a result of the intervention (0.112 of a standard deviation). This scale incorporated the students' answers to questions such as whether parents checked the child's homework, or provided motivation to them, or talked to them when needed. Moreover, the treatment also increased students' perception of their parents' level of school involvement (0.117 of a standard deviation). This perception was reflected in students' answers to questions about whether their parents contacted the school director or teachers, or whether their parents

[55]The pair-wise correlations between grades in math, language, natural science and history are, in our sample, always larger than 0.6.

[56]See Section F.1 in the Online Appendix for details on how the scales were built, as well as the psychometric properties of each of them. Appendix Table 6 present results for the individual items in each aggregate scale.

Table 7: Treatment Effects on Parental Behavior at Home

	Study habits	Academic efficiency	Family Support	Low Family Supervision	Parent School Involvement	Positive reinforcement
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Std. parent scales						
T	–0.086	0.088	–0.007	0.019	0.026	–0.057
	[0.079]	[0.064]	[0.082]	[0.064]	[0.063]	[0.080]
Observations†	1042	1090	1108	1096	1116	1098
Panel B: Std. student scales						
T	0.049	–0.005	0.112*	–0.073	0.117**	0.015
	[0.059]	[0.059]	[0.061]	[0.049]	[0.055]	[0.057]
Observations†	1726	1728	1686	1693	1700	1692

Note: Panel A and Panel B shows intention-to-treat (T) estimates on parent and student standardized scales (means are zero and standard deviations are one for the control group for all scales), respectively, and its corresponding standard error estimated using equation (1) using OLS. Outcomes are scales built with answers to surveys (see Tables F.2 and F.3 for details). All models include the baseline math grade, attendance rate and outcome scales as control variables, and classroom (randomization strata) and year and fixed effects. If baseline values of baseline math grade/attendance or baseline outcomes were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets). †Number of observations vary by column because of survey and item non-response; see Online Appendix G for details. * significant at 10%; ** significant at 5%; *** significant at 1%.

attended school meetings.

Overall, results in this section show that exposure to the text messages treatment reduced parent-school information gaps and increased student reports of parent engagement with their day-to-day school activities. Although smaller samples in the parent and student surveys make it harder to precisely estimate effects, the pattern of results is consistent with those from Bergman (2021), Bergman & Chan (2021), and Barrera-Osorio et al. (2020) who find that the additional information provided to parents increased their contact with the school.

8 Cost-Effectiveness, Willingness to Pay, Potential to Scale

8.1 Cost-effectiveness

The literature on information interventions to improve learning gains in school settings has burgeoned in recent years. Several reviews of this work now exist. For example, JPAL (2020) review the results of 23 randomized evaluations from low-, middle-, and high-income countries in which information is provided to parents about student performance (e.g., attendance, behavior, or grades) using text messages, emails, report cards, and videos. Escueta et al. (2020) find 13 field experiments where information is sent to parents about student performance through text messages and emails. Collectively, these studies show that closing knowledge gaps about a child's education often increases parental engagement with schools, student effort in school, or both, while also improving learning outcomes. Bergman (2021) is a leading example of this type of work. In his study, parents of 462 students in Los Angeles schools were randomly assigned to receive automated texts about missing assignments and grades. After four months, the text message intervention decreases the number of missed classes by 28% with a corresponding gain of 0.21 standard deviations in math grades, but no gains in English. These results are much larger than the ones we find in Chile.[57] This difference reflects an emerging fact coming from these information interventions: impacts of information interventions to improve learning in schools are high variance across different settings (Angrist, Evans, Filmer, Glennerster, Rogers & Sabarwal 2020).

Where do our estimates fit with respect to this literature? Our estimated learning gains in math (0.09 s.d.) are at the lower end of the range of effect sizes in the literature (0.09-0.19 s.d. of test scores), while our attendance gains of 1.1. percentage points fall in the middle of the range (0-2.1 percentage point gains in attendance).[58] We do find larger estimated effects for the most at-risk students in our sample of schools: for this group, the effect of the text messaging program generates grade and attendance effects at the upper end of the range of average effects in the literature.

Regarding our intervention cost, as pointed out by Bergman & Chan (2021), interventions that leverage technology to connect schools with parents on an ongoing basis are characterized by low variable cost and a once in a life-time setup cost. In their study, the variable cost per text messages was negligible, while there was a once-off fixed training cost of US$7 per student if schools did not have electronic gradebooks. In the case of Chile, the market value of sending text messages is US$0.05 per message. With an average of 6 text messages sent per month for 10 months, this adds up to $3.00 per student per year. In addition, the monthly subscription fee for a digital text messaging platform is $0.77 per student, or $7.70 per student per year. We estimate the cost of digital data entry to be $0.16 per student per year.[59] The total variable cost per student per year (in 2021 nominal prices) is there-

[57]We expected a smaller impact for our intervention, as in the United States the GPA depends on assignment submission (a directly targeted outcome in Bergman (2021)), whereas in Chile grades are based only on performance on class exams.

[58]The ranges provided here are taken from our summary of the literature in Online Appendix A.

[59]The hourly minimum wage in 2021 in Chile was approximately 2.83 dollars. Assuming that it takes an administrative staff about 5 seconds to enter the weekly attendance and grade data for each student, the total annual time allocated to data entry would be (5×40 weeks=) 200 seconds per student. Therefore, the annual cost of data entry per student amounts to 0.16 dollars ($200/60\times$ $(2.83/60)$.

fore \$10.86 per year. Given our effect sizes for math grades, the cost of a 0.01 standard deviation in math grade would be \$1.21 at market prices (10.86/9). In transitioning to this system, schools also incurred a fixed messaging platform set up cost of \$615.4 per school. Considering the average primary school size in our sample had 377 students, the fixed cost per student in the first year was \$1.63. In the first year of using a platform like Papas al Dia for parent-school communication, the cost of a 0.01 standard deviation of math grade improvement would therefore be \$1.39, with that cost falling over time.[60]

A program like Papas al Dia is cost-effective when compared to other interventions designed to improve learning outcomes. Busso et al. (2017) reviews results from 21 low-cost interventions designated to improve student learning in primary schools in Latin America and the Caribbean. Strategies include tracking, funding for materials, lesson plans, non-monetary incentives and guided technology. The authors of that study calculate the implementation cost of each intervention implemented in Colombia. The average cost per student for a 0.01 standard deviation gain in learning is US\$4.42, and the median cost is US\$2.00.[61] In terms of cost, our intervention compares favorably to these other approaches.

8.2 Willingness to Pay

In addition to being cost-effective, most parents in our study seemed willing to pay enough to cover the costs of the intervention. In our follow-up surveys, we asked both treatment and control parents to tell us whether they would be willing to pay for a text message service that provided them with four monthly messages from schools about their child's performance and behavior in school. This was a non-incentivized survey experiment in which we randomized the price at which parents were given a "take it or leave it" offer: a high price of 1,500 CLP (Chilean pesos, or 2.2 USD) per month, a medium price of 1,000 CLP (or 1.5 USD) per month, or a low price of 500 CLP (0.74 USD) per month.[62] The low price covers more than

[60]The cost of putting the experiment into the field was higher, as we had to hire a team of research assistants to visit schools, photocopy classroom books, and digitize the data.

[61]Busso et al. (2017) also provides information for 52 evaluations designed to improve student learning in secondary schools around the world. The strategies for which they find evidence of success include: i) monetary incentives to students, ii) "no excuses" models, iii) extended school day, and iv) vouchers, subsidies or scholarships for students. The weighted averages of the effect-sizes on test scores are respectively 0.16SD, 0.14SD, 0.08SD and 0.03SD. Although this study does not include intervention costs for these alternative strategies, it is likely that our text message intervention used fewer resources than any of these four programs, and therefore was cheaper on a per student basis. McEwan (2015) provides a meta-analysis of randomized experiments of school-based interventions on learning in primary schools and finds seven experiments that involve informational treatments. The mean effect size of these interventions is 0.049 (p-value=0.240). Andrabi et al. (2017) find that providing report cards to parents in Pakistan leads to a closing in informational gaps and a 0.11SD gain in student outcomes.

[62]This method of asking about willingness to pay has two important shortcomings. First, this is a hypothetical scenario. Therefore, parents have no incentive to reveal the true valuation for the service.

twice the monthly cost of sending messages.

Table 8: Parental Willingness to Pay

	[1]	[2]
Medium Price	–0.151***	–0.085
	[0.043]	[0.062]
High Price	–0.238***	–0.256***
	[0.039]	[0.059]
T x Low Price		0.030
		[0.063]
T x Medium Price		–0.095
		[0.059]
T x High Price		0.070
		[0.069]
Constant	0.706***	0.721***
	[0.264]	[0.263]
Observations	1,124	1,124

Note: Outcome is an indicator variable for whether the parent reports being willing to pay for continued text message service (4 text messages per month from the school) after the end of the year. Column [1] reports estimates of being assigned a particular randomized priced (1,500 CLP, 1,000 CLP or 500 CLP, the omitted category). Column [2] shows intention-to-treat estimates by interacting these randomized prices with the randomized treatment (equal to 1 if parents were sent text-messages and zero otherwise). All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8 uses the survey experiment to estimate parents' demand curves for the complete sample in column 1. On average, 71 percent of parents said that they were willing to pay at least the minimum amount to receive text messages from the school which generously covers the break-even costs of the intervention. In column 2 we allow each experimental group to have a different response to the randomized price by including price assignment by treatment assignment interaction terms.

Overall, the demand curve for a service like the one we offered in our intervention is downward sloped. Column 1 shows that the share of parents willing to pay for the service falls by more than 15 percentage points as the price increases from low to medium levels, and by an additional 8.7 percentage points when the price increases from a medium to a high level (the coefficient on High price is -0.238). We then analyze whether the treatment induced parents to value the text messages program differently (column 2). There is no evidence that treated parents value the information differently than control parents.

Second, we use a take-it-or-leave-it offer which gives a bound rather than an exact measure of willingness to pay.

8.3 Features of Scalability

The primary goal of this project was to evaluate an intervention that leverages *existing* school resources and practices– rather than requiring substantial additional resources or a change in school practices– to improve student outcomes. In middle-income countries, and in poor schools of high-income countries, education expenditures are already high. There are potentially large returns to adopting low-cost interventions that can make existing school expenditures more effective. Our results indicate that a text messaging intervention to improve parent-school communication can do just this.

From our experience in the field, it would be relatively straightforward for a school district to scale a Papas al Dia-like program by adopting the following three components: (1) a subscription to a text-messaging platform such as the one used in our study, possibly paid for or subsidized by parents; (2) a weekly digitization of attendance, grades, and behavior classroom books, which is already being done in some schools, but alternatively could be completed by existing administrative staff at schools; and (3) a registry of cellphone numbers for parents/guardians of students, updated at least once per year. Schools already collect contact details for parents, but contact lists would need to be digitized and shared with the digital messaging platform.

Schools in Chile have already started down the road of adopting text messaging technologies to improve communication with parents, even in the absence of national policy about such programs. When we began this study in 2014 the market for digital information platforms serving schools was nascent. In the last several years a number of companies have entered this market (e.g., one of the suppliers, Papinotas, offers various digital services to over 2,000 Chilean schools). The results from our study suggest that the expansion of these types of services in upper-primary and middle schools would likely lead to small but meaningful improvements in grades and attendance, especially for those students most at-risk of repeating grades, or dropping out, later in life. And, our positive results on spillovers to the treated group suggest that a scaled-up version of the program in which all students are treated would continue to yield positive learning and attendance gains.

9 Conclusions

We present the results of a simple, effective, and low-cost intervention that uses existing data regularly collected by schools to improve the accuracy and timeliness of information parents have about their children' attendance, grades and classroom behavior.

We showed that high-frequency text messages communicating this information to parents decreased prevailing information gaps between parents and schools, and shifted some aspects of parent-school and parent-student engagement. The intervention sustained over two school years resulted in learning and attendance gains on average, with significantly larger gains for students most at risk of poor schooling outcomes later on in life. At a broad level, our findings suggest that efforts to reduce grade retention and school dropout in later grades may be supported by early information interventions. We leave the analysis of these long-term impacts to future work.

References

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.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20140774*

Angrist, N., Bergman, P. & Matsheng, M. (2020), 'School's out: Experimental evidence on limiting learning loss using "low-tech" in a pandemic', *Available at SSRN 3735967* .

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, F. H. & Sabarwal, S. (2020), 'How to improve education outcomes most efficiently? a comparison of 150 interventions using the new learning-adjusted years of schooling metric'.

Avvisati, F., Gurgand, M., Guyon, N. & Maurin, E. (2014), 'Getting parents involved: A field experiment in deprived schools', *Review of Economic Studies* **81**(1), 57–83.

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.

Bassi, M., Busso, M. & Muñoz, J. S. (2015), 'Enrollment, graduation, and dropout rates in Latin America: Is the glass half empty or half full?', *Economía* pp. 113–156.

Bennett, M. & Bergman, P. (2021), 'Better together? social networks in truancy and the targeting of treatment', *Journal of Labor Economics* **39**(1), 1–36.

Bergman, P. (2021), 'Parent-child information frictions and human capital investment: Evidence from a field experiment', *Journal of Political Economy* **129**(1).

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* pp. 1118–9837R1.

Bergman, P., Lasky-Fink, J. & Rogers, T. (2020), 'Simplification and defaults affect adoption and impact of technology, but decision makers do not realize it', *Organizational Behavior and Human Decision Processes* **158**, 66–79.

Bettinger, E., Cunha, N., Lichand, G. & Madeira, R. (2021), 'Are the effects of informational interventions driven by salience?', *University of Zurich, Department of Economics, Working Paper* **350**.

Bursztyn, L. & Coffman, L. C. (2012), 'The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas', *Journal of Political Economy* **120**(3), 359–397.

Busso, M., Cristia, J., Hincapié, D., Messina, J. & Ripani, L. (2017), *Learning Better: Public policy for skills development*, Inter-American Development Bank.

Castleman, B. L. & Page, L. C. (2015), 'Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates?', *Journal of Economic Behavior & Organization* **115**, 144–160.

Cattaneo, M. D., Jansson, M. & Ma, X. (2018), 'Manipulation testing based on density discontinuity', *The Stata Journal* **18**(1), 234–261.

De Walque, D. & Valente, C. (2018), *Incentivizing school attendance in the presence of parent-child information frictions*, The World Bank.

Dinkelman, T. & Martínez A, C. (2014), 'Investing in schooling in Chile: The role of information about financial aid for higher education', *Review of Economics and Statistics* **96**(2), 244–257.

Dizon-Ross, R. (2019), 'Parents' beliefs about their children's academic ability: Implications for educational investments', *American Economic Review* **109**(8), 2728–65.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20171172*

Escueta, M., Nickow, A. J., Oreopoulos, P. & Quan, V. (2020), 'Upgrading education with technology: Insights from experimental research', *Journal of Economic Literature* **58**(4), 897–996.
URL: *https://www.aeaweb.org/articles?id=10.1257/jel.20191507*

Gallego, F. A., Malamud, O. & Pop-Eleches, C. (2020), 'Parental monitoring and children's internet use: The role of information, control, and cues', *Journal of Public Economics* **188**, 104208.

JPAL (2020), "Improving learning outcomes through providing information to students and parents", *J-PAL Policy Insights*, Abdul Latif Jameel Poverty Action Lab (JPAL).
URL: *https://www.povertyactionlab.org/policy-insight/providing-information-students-and-parents-improve-learning-outcomes*

Kraft, M. A. & Dougherty, S. M. (2013), 'The effect of teacher–family communication on student engagement: Evidence from a randomized field experiment', *Journal of Research on Educational Effectiveness* **6**(3), 199–222.

Kraft, M. A. & Rogers, T. (2015), 'The underutilized potential of teacher-to-parent communication: Evidence from a field experiment', *Economics of Education Review* **47**, 49–63.

Lim, M., Tormshak, E. A. S. & Dishion, T. J. (2005), 'A one-session intervention for parents of young adolescents: Videotape modeling and motivational group discussion', *Journal of Emotional and Behavioral Disorders* **13**(4), 194–199.

Manacorda, M. (2012), 'The cost of grade retention', *The Review of Economics and Statistics* **94**(2), 596–606.

McEwan, P. J. (2015), 'Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments', *Review of Educational Research* **85**(3), 353–394.

Mizala, A., Romaguera, P. & Urquiola, M. (2007), 'Socioeconomic status or noise? trade-offs in the generation of school quality information', *Journal of development economics* **84**(1), 61–75.

OECD (2022), 'Secondary graduation rates (indicators)'. Accessed: 2022-06-06.
URL: *https://data.oecd.org/students/secondary-graduation-rate.htm*

Rogers, T. & Feller, A. (2018), 'Reducing student absences at scale by targeting parents' misbeliefs', *Nature Human Behaviour* **2**(5), 335–342.

Tincani, M. M. (2018), Heterogeneous peer effects in the classroom, Technical report, Working Paper.

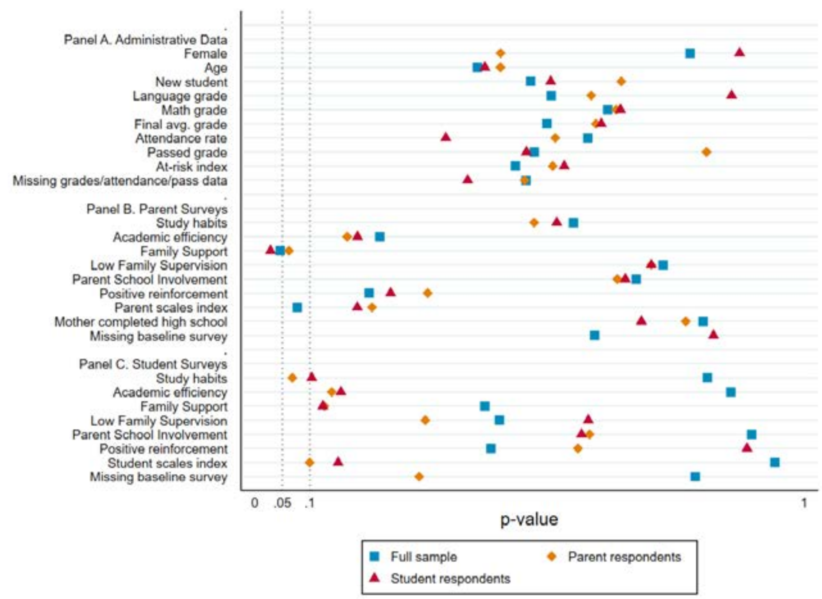
Wedenoja, L. (2017), 'The dynamics of high school dropout', *unpublished* .
URL: *https://sites.google.com/site/leighwedenoja/research*

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.

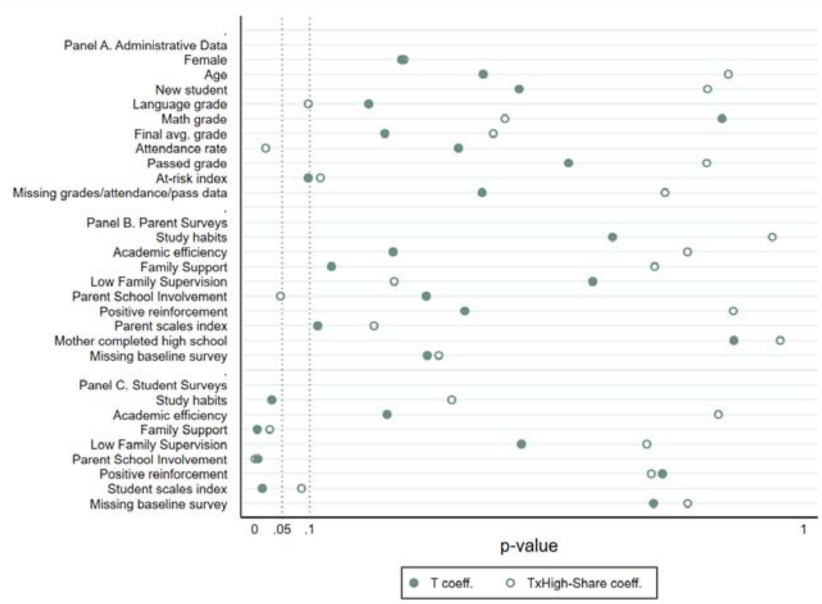
Appendix Figures

Figure 1: Balance in alternative samples and specification

(a) Alternative samples



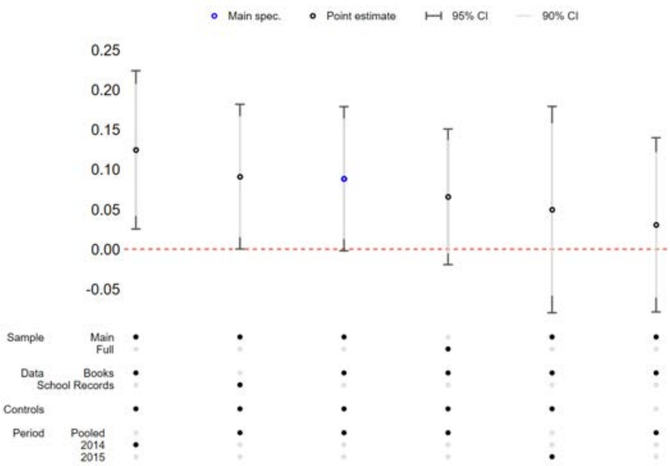
(b) High-share specification



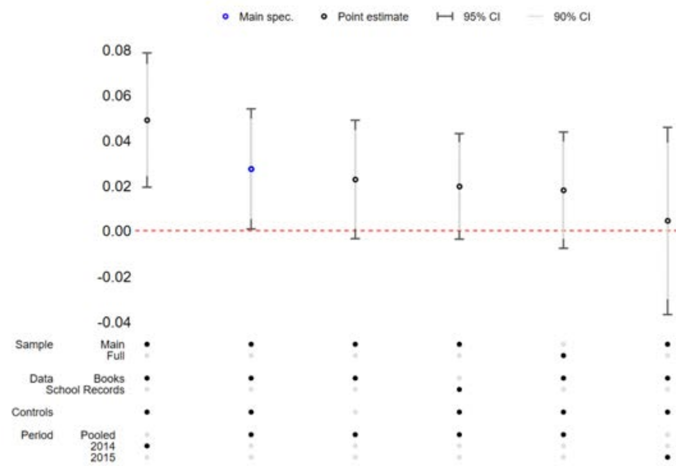
Note: Panel A plots the p-value on the treatment coefficient in a regression using each baseline characteristic as the dependent variable for alternative samples (full sample, surveys' parent and student respondents). Panel B plots p-values on the treatment coefficient and on the interaction between treatment and high-share classrooms in regressions using each baseline characteristic as the dependent variable. All regressions include classroom fixed effects and robust standard errors are clustered at this level. Observable variables correspond to 2013 except for new student variable that refers to 2014.

Figure 2: Robustness

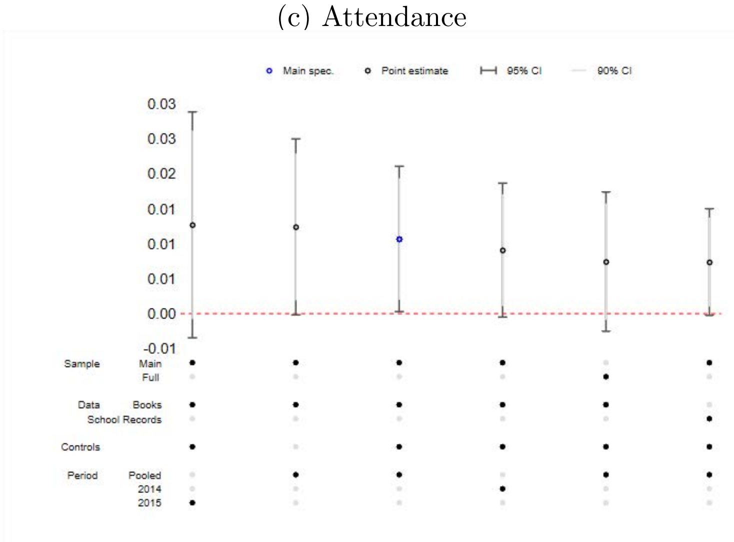
(a) Standardized math grade



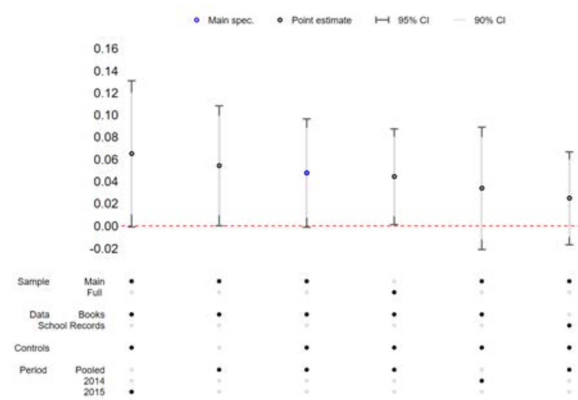
(b) Math grade > 4



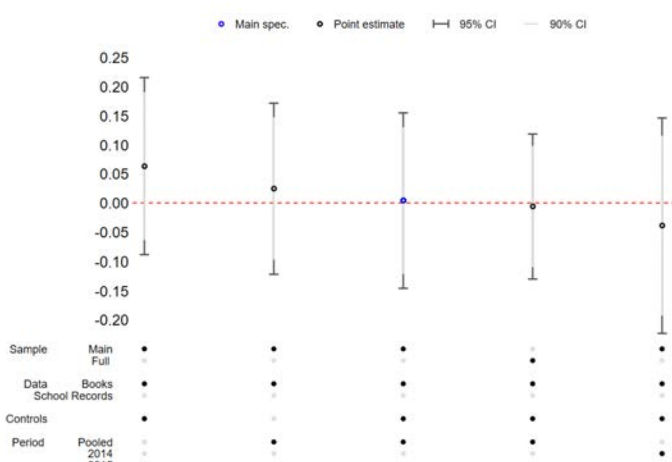
(c) Attendance



(d) High attendance



(e) Standardized # negative behavioral notes



Note: Figure plots the treatment coefficients (and 90% and 95% confidence intervals) for each panel outcome using different specifications (with and without baseline controls), different samples (with and without the students who leave the sample due to being in grade 8 in the first year, or being in one school that left our sample in the second year; using the pooled sample versus separating the midline and endline samples), and different data sources for outcomes (using administrative data from the national ministry or administrative data from the school records collected by our research team). Each combination is represented by black/white dots in the bottom of each subfigure.

Appendix Tables

Table 1: Retention and drop-out

	Retention	Drop-out	Retention	Drop-out
	[1]	[2]	[3]	[4]
GPA_{t-1}	–0.034***	–0.005***		
	[0.000]	[0.000]		
GPA_{t-2}	–0.019***	–0.001***		
	[0.000]	[0.000]		
GPA_{t-3}	–0.039***	0.000		
	[0.000]	[0.000]		
$Attendance_{t-1}$	–0.003***	–0.001***		
	[0.000]	[0.000]		
$Attendance_{t-2}$	–0.001***	–0.001***		
	[0.000]	[0.000]		
$Attendance_{t-3}$	–0.000	–0.001***		
	[0.000]	[0.000]		
$At-risk\,index_{t-1}$			0.076***	0.027***
			[0.000]	[0.000]
$At-risk\,index_{t-2}$			0.037***	0.017***
			[0.000]	[0.000]
$At-risk\,index_{t-3}$			0.037***	0.007***
			[0.000]	[0.000]
Observations	6,594,877	6,594,877	6,594,877	6,594,877
Adjusted-R2	0.116	0.0970	0.0944	0.0522

Note: Table shows estimates of a linear probability model with retention or drop-out in year t as dependent variable. Columns 1-2 show standardized GPA attendance $t-k$ years ago ($k = 1, 2, 3$) estimate coefficients. Columns 3-4 estimate the same lags for an at-risk index. At-risk index is the negative of a simple average of standardized attendance and GPA. Based on public data for primary and secondary education level for the period 2002-2020 from the Ministry of Education of Chile. We restrict the sample to educational trajectories of students who were in grades 8-12 between 2006 and 2013 and that ever attended any school in the Santiago metropolitan region. Grades 1-3 are excluded. All models control for student's sex and include municipality fixed effects. Standard errors clustered at the student level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Compliance in different samples

	All	Attendance	Behavior	Grades	General
	[1]	[2]	[3]	[4]	[5]
Panel A: 2014					
Text messages sent	29.600***	21.079***	4.509***	4.021***	–0.008
	[0.365]	[0.260]	[0.074]	[0.052]	[0.029]
Text messages received	20.080***	14.289***	3.085***	2.748***	–0.042
	[0.649]	[0.402]	[0.120]	[0.093]	[0.083]
Observations	1063	1063	1063	1063	1063
Panel B: 2015					
Text messages sent	60.403***	40.129***	9.236***	11.111***	–0.073
	[1.366]	[0.886]	[0.159]	[0.259]	[0.158]
Text messages received	33.464***	21.452***	6.125***	6.144***	–0.257
	[1.284]	[0.720]	[0.206]	[0.208]	[0.250]
Observations	948	948	948	948	948
Panel C: Full Sample					
Text messages sent	42.178***	28.879***	6.456***	6.890***	–0.047
	[0.723]	[0.452]	[0.097]	[0.149]	[0.067]
Text messages received	25.463***	17.162***	4.344***	4.127***	–0.170
	[0.722]	[0.422]	[0.119]	[0.124]	[0.104]
Observations	2439	2439	2439	2439	2439
Panel D: Parent Surveys 2015					
Declares to have received text messages	0.359***	0.523***	0.431***	0.443***	-
	[0.049]	[0.042]	[0.042]	[0.047]	-
Observations	549	565	561	567	-

Note: Panel A uses the 2014 data of the intervention. Panel B uses the 2015 data of the intervention. Panel C analyzes compliance in the full sample. Panel D uses 2015 parents' surveys data. Text messages sent/received refers to the cumulative number of text messages sent to/received by student's parents. For Panels A-C columns [2]-[5] report the T_{icjg} coefficient of equation (1) with the annual number of each type of text message as the dependent variable. Column [1] adds all types of text messages. For Panel D columns [1]-[4] report the T_{icjg} of equation (1) using each column parent's self-declared text messages' reception as the dependent variable. Parents answer on a four-value scale the frequency in which they have received each type of text message ("never or almost never" to "always or almost always") in the last month. Outcomes are indicator variables equal to one if parent answer value 4 and zero otherwise. Column [1] outcome equals one if at least one of the attendance, grades and behavior text messages outcomes equals one. Attendance, grades, and classroom behavior text messages were sent only to the treatment group. General text messages were sent to all treatment and control individuals. All models include the baseline math grade and attendance rate as control variables. If baseline values are missing, we impute them using the classroom-level mean and flag these observations in the regression. Regressions additionally include year and classroom fixed effects and standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Local Average Treatment Effects

	Standardized math grade	Math grade >4.0	Attendance rate	Cumulative attendance >85%	Standardized # negative beh. notes
	[1]	[2]	[3]	[4]	[5]
Panel A: LATE					
D	0.106*	0.033**	0.012**	0.055*	0.005
	[0.054]	[0.016]	[0.006]	[0.028]	[0.090]
Panel B: Heterogeneity					
D	0.108**	0.032**	0.012**	0.054*	–0.027
	[0.053]	[0.016]	[0.006]	[0.028]	[0.080]
D x at-risk index	0.172**	0.030	0.016*	0.084***	–0.256**
	[0.085]	[0.022]	[0.008]	[0.031]	[0.115]
Observations	2011	2011	2011	2011	2011
Control mean	0.00	0.934	0.877	0.728	0.00

Note: Panel A shows estimates of the local average treatment effects (LATE) shown on each column for each outcome. Let D_{icjg} be an indicator variable equal to one for those treated students whose parents received at least one text message with information on each specific outcome (i.e., compliers). D_{itcjg} –instead of T_{icjg}– is included in equation (1) which we instrument in a first stage with the randomized treatment variable T_{icjg}. Panel B adds the interaction with the student-level at risk index. At-risk index is a simple average of standardized baseline attendance, math grades and negative behavioral notes. All models include the baseline math grade, attendance rate as control variables, classroom (randomization strata) and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Models in Panel B additionally include the at-risk index variable as control. Columns 1 and 5 report results on outcomes that were standardized so that mean among the control students is zero and the standard deviation is one. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Treatment Effects Over the Week (Weekly Fade Out)

	Daily Attendance
T x Monday	0.015**
	[0.007]
T x Tuesday	0.017**
	[0.007]
T x Wednesday	0.014*
	[0.007]
T x Thursday	0.005
	[0.006]
T x Friday	0.006
	[0.007]
Observations	222827
p-value of equal coeff.	0.037
p-value of TxMonday = TxFriday†	0.065

Note: Table shows intention-to-treat estimates (T) by day of the week estimated using OLS. Attendance outcome is measured at a daily basis. T refers to the randomized treatment (equal to 1 if parents were sent text-messages and zero otherwise) and is interacted with each day-of-the-week indicator variables. All models include the day-of-the-week indicator variables as controls, baseline math grade, attendance rate as control variables, classroom (randomization strata), and month x year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets). † One-sided test against the alternative that TxFriday > TxMonday. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Treatment Effects Over Time

	Standard. math grade	Attendance rate	Standard. # negative beh. notes
	[1]	[2]	[3]
T x months 1–3	0.114**	0.007	–0.014
	[0.057]	[0.006]	[0.073]
T x months 4–6	0.053	0.009	–0.024
	[0.054]	[0.007]	[0.044]
T x months 7–9	0.068	0.004	0.059
	[0.071]	[0.007]	[0.076]
T x months 10–12	0.109**	0.016*	0.042
	[0.054]	[0.009]	[0.050]
T x months 13–17	0.054	0.018*	0.015
	[0.051]	[0.010]	[0.052]
Observations	10,391	15,912	15,568
p-value of equal coeff.	0.766	0.751	0.584

Note: Table reports intention-to-treat (T) estimates for each group-of-months estimated using OLS. Outcomes are measured at a monthly basis. T refers to the randomized treatment (equal to 1 if parents were sent text-messages and zero otherwise) and is interacted with each group-of-months indicator variables. All models include the group-of-months indicator variables as controls, baseline math grade, attendance rate as control variables, classroom (randomization strata), month and year fixed effects. If baseline values of baseline math grade/attendance were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Treatment Effects on Parental Behavior at Home: Indicators

	Organized for school work	Understand difficult content	Parents show pride	Went to school alone	Parents contacted teacher	Parents congratulated student
	[1]	[2]	[3]	[4]	[5]	[6]
Panel A: Parent scales						
T	–0.063*	0.018	0.004	0.021	0.010	–0.027
	[0.032]	[0.028]	[0.020]	[0.038]	[0.031]	[0.023]
Control mean	0.769	0.871	0.943	0.364	0.294	0.927
Observations†	1125	1112	1164	1161	1168	1158
Panel B: Student scales						
T	0.008	0.010	0.031	–0.020	0.049*	0.012
	[0.030]	[0.021]	[0.029]	[0.025]	[0.026]	[0.025]
Control mean	0.656	0.816	0.720	0.466	0.334	0.746
Observations†	1787	1781	1761	1784	1772	1772

Note: Table shows intention-to-treat effects estimates from equation (1) shown on each column for each outcome. Coefficients were estimated using OLS. T refers to the randomized individual-level treatment (equal to 1 if parents were sent text-messages and zero otherwise). Outcomes are behavior indicators built with answers to surveys (see Tables F.2 and F.3 for details). For each scale, we take the item with the largest loading factor and build an indicator variable that takes value 1 when student/guardian answer 3 or 4 in the four scale. Items are (student versions): 'I organize well my time to do my school work', I am sure that I can understand the hardest things, My parents or guardians showed that they were proud of me, I went alone to school, My parents or guardians contacted teacher through e-mail, My parents or guardians congratulated me for my effort. Panel A shows results for scales built with answers parents gave to survey questions. Panel B shows results for scales built with answers students gave to survey questions. All models include the baseline math grade, attendance rate and outcome scales, classroom (randomization strata), and year fixed effects. If baseline values of baseline math grade/attendance or baseline outcomes were missing, we imputed them using the classroom-level mean and added an indicator variable for these imputed observations. Standard errors are clustered at the classroom level (shown in brackets).† Number of observations vary by column because of survey and item non-response. * significant at 10%; ** significant at 5%; *** significant at 1%.