
Leveraging Parents through Low-Cost Technology

The Impact of High-Frequency Information on Student Achievement

Peter Bergman
Eric W. Chan

ABSTRACT


We partnered a low-cost communication technology with school information systems to automate the gathering and provision of information on students' academic progress to parents of middle and high school students. We sent weekly automated alerts to parents about their child's missed assignments, grades, and class absences. The alerts reduced course failures by 27 percent, increased class attendance by 12 percent, and increased student retention, though there was no impact on state test scores. There were larger effects for below-median GPA students and high school students. More than 32,000 text messages were sent at a variable cost of \$63.

Peter Bergman is an assistant professor of economics and education at Teachers College, Columbia University (bergman@tc.columbia.edu). Eric Chan is an assistant professor of statistics and public policy at Babson College. The authors thank Kanawha County Schools, as well as Spencer Kier, Alex Farivar, Jeremy Lupoli, Sam Elhag, and Zach Posner. They also thank Raj Darolia, Philipp Lergetporer, and Susanna Loeb for their comments, as well as those of seminar participants at the Wisconsin Institute for Research on Poverty, IZA, and CESifo. This research is funded by the Smith Richardson Foundation and the Teachers College Provost's Investment fund and received approval from the Institutional Research Board at Teachers College, Columbia University. The experiment described is pre-registered on AEA RCT Registry (registration #AEARCTR-0001227). Any errors are those of the authors. Bergman has previously received compensation from the learning management system company to advise them with respect to their school-to-parent information technologies. Student-level data for this study are available only with permission from Kanawha County Schools. Exploratory analysis also includes American Community Survey data, which are publicly available.

[Submitted November 2018; accepted February 2019]; doi:10.3368/jhr.56.1.1118-9837R1

JEL Classification: I20, I21, I24, and I28

ISSN 0022-166X E-ISSN 1548-8004 © 2021 by the Board of Regents of the University of Wisconsin System

 Supplementary materials are freely available online at: <http://uwpress.wisc.edu/journals/journals/jhr-supplementary.html>

I. Introduction

Families are both one of the greatest sources of inequality and a powerful determinant of academic achievement (see Coleman et al. 1966; Heckman 2006; Cunha and Heckman 2007; Todd and Wolpin 2007). While leveraging families has the potential to improve child outcomes, many programs that do so focus on skills-based interventions, can be difficult to scale due to high costs, and often concentrate on families with young children (Belfield et al. 2006; Olds 2006; Nores and Barnett 2010; Heckman et al. 2010; Avvisati et al. 2013; Duncan and Magnuson 2013; Gertler et al. 2014; Mayer et al. 2015; Doss et al. 2018; York, Loeb, and Doss 2018).

However, for older children, recent research suggests it may be possible to leverage parents by resolving certain intrahousehold information frictions around their child's effort in school.¹ For instance, Bursztyn and Coffman (2012) conducted a lab experiment that shows families prefer cash transfers that are *conditional* on their child's school attendance over unconditional cash transfers because this helps them monitor their child. When parents are offered a chance to receive information about their child's attendance via text message, they no longer are willing to pay for conditionality. Bergman (2014) conducted an experiment that sent information to parents about their child's missing assignments to estimate a persuasion game (Dye 1985; Shin 1994) in which children strategically disclose information about their effort to parents, and parents incur costs to monitor this information. He found that, for high school students, the additional information reduces parents' upwardly biased beliefs about their child's effort and makes it easier for parents to monitor their child.² There is evidence that students' effort and achievement improve as a result.

The idea that providing information to parents about their child's effort can improve outcomes is tantalizing. There is a lack of low-cost interventions that can successfully improve education outcomes for children during middle and high school (Cullen et al. 2013).³ While there is promise that providing high-frequency information about students' progress could impact outcomes, the current evidence comes from small, labor-intensive interventions, some of which do not look directly at measures of student achievement and may be difficult to scale. For instance, Kraft and Rogers (2014) show that personalized messages, written individually by teachers and sent by research staff to a sample of 435 parents, helped retain students in a summer credit-recovery program. Bergman (2014) sent bimonthly text messages—typed by hand—and phone calls about students' missed assignments and grades to a sample of 279 parents. In theory, placing student information online could help resolve these information problems, but Bergman (2016) finds that parent adoption and usage of this technology is low, especially in schools serving lower-income and lower-achieving students, which could exacerbate socioeconomic gaps in student achievement.

1. Parents also exhibit information problems about their child's ability, attendance, and the education production function (Bonilla et al. 2005; Bursztyn and Coffman 2012; Cunha, Elo, and Culhane 2013; Bergman 2014; Rogers and Feller 2016; Dizon-Ross 2019; Kinsler and Pavan 2016; Andrabi, Das, and Khwaja 2017). Fryer (2011) also finds that students may not accurately assess the education production function as well.

2. The experiment became contaminated for middle school students when the school asked an employee to call all of these students' parents about their missed assignments.

3. There are also a number of low-cost interventions that focus on the transition from high school to college, which have improved college enrollment outcomes (such as Bettinger et al. 2012; Castleman and Page 2015; Hoxby and Turner 2013; Carrell and Sacerdote 2017).

We develop and test a low-cost technology that synchronizes with student information systems and teacher gradebooks to push high-frequency information to parents about their child's absences, missed assignments, and low grades via automated text messages.⁴ The intervention automates sending out three types of alerts. First, an absence alert was sent weekly detailing the number of classes a child missed for each course in the last week. This by-class alert contrasts with how districts typically report absences to parents, which are usually reported in terms of full-day absences.⁵ Similarly, if a student missed at least one assignment over the course of a week, a weekly alert was sent stating the number of assignments missed in each class during the past week. Finally, a low-grade alert was sent once per month if a child had a class grade average below 70 percent at the end of the month. Messages were randomly assigned to be delivered to either the mother or the father.

We conducted a field experiment to evaluate the academic and behavioral effects of this information technology in 22 middle and high schools. In the first year of the intervention, we sent 32,472 messages to treatment group families, or an average of 52 messages per treated family. This increased the likelihood parents were contacted by schools at least once per month by 19 percentage points. This high-frequency contact contrasts with the existing amount of contact between schools and parents, which varies widely. Our surveys indicated that nearly 50 percent of parents were contacted less than one time in three months by the school about their child's academic progress, but roughly one-quarter of parents reported hearing from their school more than once per month. We also find that parents tend to overestimate their child's grades, on average, and to underestimate their child's missed assignments. Parents are more accurate about the former than the latter.

We find that, as a result of this additional contact, there was a substantial decrease in the number of courses students failed. In the first year, students failed one course, on average, and the text-message intervention reduced this by nearly 27 percent. Treatment group students also attended 12 percent more classes, and district retention increased by 1.5 percentage points. We do not find any improvements in state math and reading test scores. However, statewide exams had no stakes for students, and students spent roughly 100 minutes less time than the test provider expected for students to complete the exams. The district subsequently discontinued using these standardized tests.⁶ In contrast, we do find significant, 0.10 standard-deviation increases for in-class exam scores.⁷

Most of the positive impacts were driven by two subgroups we prespecified in our analysis plan: students with below-median grade point averages (GPAs) and high school

4. Text messages have been used in several education interventions (Kraft and Rogers 2014; Bergman 2014; Castleman and Page 2015, 2016; Page, Castleman, and Meyer 2016; Oreopoulos and Petronijevic 2018; Castleman and Page 2017).

5. The definition of a full day can vary from district to district. In this district, it constitutes 70 percent of the school day.

6. The state superintendent's commission expressed concerns that the exam is not "an accurate gauge of student achievement" and "doesn't give much reason for students to take it seriously," as quoted in the *Charleston Gazette Mail* (Quinn 2016a).

7. Several recent studies have documented substantial increases in student performance in response to the stakes of the exam (List, Gneezy, and Sadoff 2017) and student effort (Hitt 2015; Zamarro, Hitt, and Mendez 2016).

students. These groups experienced larger impacts on grades, GPA, missed assignments, course failures, and retention, while middle-school students show no significant, positive (or negative) effects. We find that the positive effects for high school students and students with below-median GPAs persisted into the second year of the intervention. We do not find a differential effect of alerting mothers versus fathers.

Our work makes several contributions to the literature. First, we show that pushing high-frequency information to parents can significantly improve student performance in middle and high schools. One novel aspect of this information, which had not been previously tested, is that we tracked and sent information to parents about every class their child had missed (as opposed to full-day absences). We show that this information is particularly important because students are 50 percent more likely to miss an individual class than a full day of school. We are among the first to show that class absences occur at much higher rates than full-day absences (see Whitney and Liu 2017).

Second, we test a unique, automated technology. The promise of automation is that, relative to other interventions, communicating with parents via automated alerts is extremely low cost, even at scale. Previous work has involved asking teachers to write tailored content about students or has used research assistants to gather and provide information to families. Our intervention allows schools to leverage their data system “as is.” Moreover, sending information via text message is affordable. Despite sending more than 32,000 text messages, the total cost was approximately \$63. If schools have no existing gradebook system, the platform and personnel training cost an additional \$7 dollars per student. This low marginal cost of the collection and provision of information implies that automated messaging has a high potential to impact student achievement at scale. Rogers and Feller (2016) also demonstrated a successful, scalable attendance intervention by mailer, but they find (as they expected) that the attendance gains, equal to an additional day of school, were not large enough to improve student performance. We complement their research by demonstrating the value of targeting students’ *by-class* absences.

Lastly, previous research in areas such as cash transfers suggests that intrahousehold divisions of labor and bargaining power could lead to heterogeneous effects of education interventions if information does not flow perfectly between parents (Doepke and Tertilt 2011; Duflo 2012; Yoong, Rabinovich, and Diepeveen 2012; Akresh, De Walque, and Kazianga 2016). To assess this possibility, we randomized whether a student’s mother or father received the intervention.

Related to our paper is ongoing work by Berlinski et al. (2016), who conducted a texting intervention in eight elementary schools in Chile. They send information to parents about their child’s math test scores, math grades, and attendance. These data are gathered from schools and entered by their research team into a digital platform, which is then used to send text messages to parents. In contrast, we automated this process by scraping data that is frequently entered into district student information systems, such as grades, attendance, and missed assignments.

Providing high-frequency information to parents about their child’s academic progress is a potentially important tool in the set of effective interventions for older children. In contrast to the intervention tested here, this set of tools often includes high-touch interventions, such as intensive tutoring (Bloom 1984; Cook et al. 2015; Fryer 2017) or one-on-one advising (Oreopoulos and Petronijevic 2018), but their scale may be limited

by high costs.⁸ However, despite the low cost and scalability of our intervention, it is not a panacea. Automation alone also does not address other aspects of scalability, such as adoption by parents and usage of the system by teachers. In Washington, DC, Bergman and Rogers (2017) randomized whether approximately 7,000 parents can enroll in this technology via an opt-in process or an opt-out process. Takeup for the opt-out process is above 95 percent, with similar effects on outcomes as in this study, and this group is more likely to opt into the technology in the future. Another potential constraint is whether teachers supply information to the gradebook system in a timely fashion. In this study teachers were blinded to the intervention assignment and we discerned no impacts of the intervention on teacher logins into the system. Some school districts contractually obligate teachers to update grades regularly (for instance, weekly). Our research suggests this type of policy could be an important—and potentially monitored—input for improving education outcomes.

The rest of the paper proceeds as follows. Section II describes the background and the experimental design. Section III describes the data collection process and outcome variables. Section IV presents the experimental design and the empirical specifications. Section V shows our results, and Section VI concludes.

II. Background and Experimental Design

A. Background

The experiment took place in 22 middle and high schools during the 2015–2016 school year in Kanawha County Schools (KCS), West Virginia. We were subsequently able to extend the intervention through the 2016–2017 school year. As a state, West Virginia ranks last in bachelor degree attainment and 49th in median household income among U.S. states and the District of Columbia. It is the only state where less than 20 percent of adults over 25 years of age have a bachelor's degree, and households have an overall median income of \$42,019.⁹ The state population is 93 percent white, 4 percent African-American, 1 percent Asian, and 2 percent Hispanic or Latino as of 2015. Data from the 2015 NAEP showed that 73 percent of West Virginia students were eligible for free or reduced-priced lunch. Students also scored significantly below the national average on all National Assessment of Educational Progress (NAEP) subjects tested, with the exception of fourth grade science, which was in line with national averages.¹⁰

Kanawha County Schools is the largest school district in West Virginia, with more than 28,000 enrolled students as of 2016. The district's four-year graduation rate is 71 percent, and standardized test scores are similar to statewide proficiency rates in

8. Cook et al. (2015) designed a novel, more affordable high-intensity tutoring intervention for male ninth and tenth grade students, which reduced course failures by 50 percent and increased math test scores by 0.30 standard deviations. The intervention, however, is costly at several thousand dollars per student. Alternatively, Oreopoulos and Petronijevic (2018) compare a one-on-one student coaching program to both a one-time online exercise and a text-message campaign and find that only the high-touch program impacts achievement.

9. American Community Survey one-year estimates and rankings by state can be found at <https://www.census.gov/acs/www/data/data-tables-and-tools/ranking-tables/> (accessed March 25, 2020).

10. NAEP Results by state can be found at <https://www.nationsreportcard.gov/profiles/stateprofile/overview/WV> (accessed March 25, 2020).

2016. In the school year prior to the study, 2014–2015, 44 percent of students received proficient-or-better scores in reading, and 29 percent received proficient-or-better scores in math. At the state level, 45 percent of students were proficient or better in reading, and 27 percent were proficient in math. 83 percent of district students are identified as white, and 12 percent are identified as Black. 79 percent of students receive free or reduced-priced lunch compared to 71 percent statewide.¹¹

Like much of the state, the district has a gradebook system for teachers. Teachers record by-class attendance and mark missed assignments and grades using this web-based platform. Teachers are obligated to mark attendance, but the only obligation to update the gradebook is every six to nine weeks, which correlates to the dissemination of report cards. We worked with the Learning Management System (LMS) provider of this gradebook to design a tool that automatically drew students' missed assignments for each class, their percent grade by class, and their class-level absences from the gradebook. The tool synchronized with the student information system to pull in parents' contact information. This allowed us to automatically pair contact information with information on academic progress from the gradebook and then push it out to families using a text-messaging API developed by Twilio. These text messages form our parent-alert system. Each of the text messages was designed to be a consistent weekly or monthly update to the parents of students who had at least one absence or missing assignment during the week or who had a low class average grade over the course of a month.

The gradebook application also has a "parent portal," which is a website that parents can log into to view their child's grades and missed assignments. All parents in the study could access the parent portal, and any parent could turn on our alerts by logging into the portal and turning on the alert feature. As we discuss further below, only 2 percent of parents in the control group received any alert. Bergman (2016) finds that, in general, very few parents ever use the parent portal, and we find this is true in KCS as well. Roughly one-third of parents had ever logged in to view their child's grades. Moreover, usage of the parent portal tends to be higher for higher-income families and families with higher-performing students.

We tested three types of parent alerts: Missed assignment alerts, by-class attendance alerts, and low-grade alerts. The text of the alerts are as follows, with automatically inserted data in brackets:

Missed Assignment Alert: "Parent Alert: [Student Name] has [X] missing assignment(s) in [Class Name]. For more information, log in to [domain]"

By-Class Attendance Alert: "Parent Alert: [Student Name] has [X] absence(s) in [Class Name]. For more information, log in to [domain]"

Low Class Average Alert: "Parent Alert: [Student Name] has a [X] percent average in [Class Name]. For more information, log in to [domain]"

If a child missed at least one assignment during a week for any course, the parent received a text-message alert reporting the number of assignments their child was missing for each course during the past week. The missing assignment message was scheduled for each Monday. These assignments included homework, classwork, projects, essays, missing exams, tests, and quizzes. On Wednesdays, parents received an alert for any class their child had missed the previous week. Finally, on the last Friday of each month, parents

11. These summary statistics come from the state education website, which can be found <https://zoomwv.k12.wv.us/Dashboard/portalHome.jsp> (accessed March 25, 2020).

received an alert if their child had a cumulative average below 70 percent in any course during the current marking period. Each alert was sent at 4:00 p.m. local time. The text messages also included a link to the website domain of the parent portal, where the parent could obtain specific information on class assignments and absences if necessary.

These alerts targeted lower-performing students. We hypothesized that impacts would be greatest for students with lower GPAs. We believed these students would have less incentive to tell their parents about their academic performance (Bergman 2014). To the extent that additional information makes it easier for parents to track their child's performance, the intervention might increase the accuracy of their beliefs about their child's performance and facilitate parents' ability to induce more effort from their child to do well in school. We explore these mechanisms in our endline survey results and stratified the randomization by baseline GPA to explore heterogeneous effects.

The rest of this section and Section III describe recruitment, the field experiment, and data collection. Figure 1 shows the timeline from the consent process through the first year of the experiment. Baseline data were collected from June to July 2015. We obtained demographic and enrollment data for the 2014–2015 school year from KCS along with contact and address information. Consent letters were sent out beginning August 2015 during the beginning of the school year. Calls requesting verbal consent were completed in September. Randomization into treatment and control was completed in early October 2015. For parents who were selected into treatment, introductory text messages were sent late that same month. Included in the texts was the option to stop at any point by replying “stop” or any equivalent variation.¹² Over the course of the study, nine parents or guardians requested the messages stop.¹³ The intervention ran from the end of October 2015 through the end of May when the school year was expected to conclude. Officially, the academic school year ended in early June, but varied slightly based on weather-induced make-up days at each school. After the end of the school year, we collected endline survey data both by phone and by mail as described below.

At the end of the 2015–2016 year, we asked KCS whether we could continue the intervention for a second year. KCS agreed to continue the intervention, but limited our data collection to outcomes originating from the gradebook.

B. Recruitment and Sample Selection

The initial sample began with approximately 14,000 total students who were enrolled in Grades 5–11 during the 2014–2015 school year. Recruitment was at the household level. A number of these students lived in the same households, so the final sample frame was 10,400 households.

During the summer of 2015, one consent letter was sent to each household in the sample frame. This letter was specifically addressed to one randomly selected parent or guardian when contact information was available for more than one parent in the data provided by the district. The letter contained the name of a randomly selected student living in the household, and this student would be the subject of the intervention.¹⁴

12. We manually tracked replies to ensure the service was shut off when requested.

13. These parents were included as “treated” families in all analyses.

14. Students were in Grades 5–11 the previous year and were expected to be in Grades 6–12 during the school year of the study.

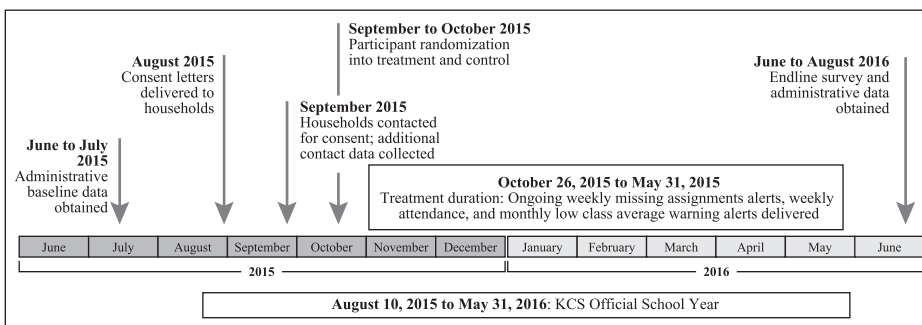


Figure 1
Field Experiment Timeline

Notes: This figure shows the timeline of the project, which began during the summer of 2015 and lasted through the summer of 2016, when data collection ended formally. The district allowed continuation of the treatment for a second school year in 2016–2017, but only district administrative data were obtained.

Trained interviewers followed up the letter with a phone call to each selected parent to confirm their participation and contact information. This was required by the district and our IRB Institutional Review Board. We then asked their language preference and preferred modes of contact: text message or phone calls. As a result, the parent or guardian of 1,137 students consented to the study and provided their contact information for inclusion as a participant.¹⁵ Though it deviated from our original design, to simplify our intervention and to save on costs, we chose to implement a text-message-only intervention. Those who could only be contacted by a landline phone or did not wish to be contacted by text did not receive the intervention even if they were randomized into treatment.¹⁶ Only 4 percent of parents could only be contacted by a landline or did not wish to be contacted by text.

We examine the possibility of self-selection into the study. Consent documents sent to parents state that the “new program uses mobile devices to communicate specific academic information with parents...via telephone and text messages.” The information provided to parents at the time of consent may appeal to parents who are more engaged in their children’s education and/or parents who are less able to keep track of their children’s academic progress. It may also appeal to those who are more technologically inclined. Table 1, Panel B, presents baseline summary statistics and a

15. Of the 14,000 students in middle and high schools in the district, many were from the same households. This decreased our sample to 10,400 households at the time. These parents were called a maximum of three attempts. Of the 10,400 phone numbers, approximately 60 percent of households never picked up the phone. About 11 percent were bad numbers that either did not work or led to the incorrect household. Another 5 percent declined to speak with the consultant. As a result, 24 percent of households both answered the phone and were willing to continue speaking with the consultant. Out of the parents willing to speak to the consultant, about 43 percent proceeded to give verbal consent to participate. A small number of consenting parents preferred phone calls over text messages, so this resulted in a final sample of about 10 percent of the 10,400 sample who were willing to accept a text-message-based alerts. Of these participants, 96 percent of the treatment and control groups preferred to receive text messages.

16. No families are dropped from the analysis, however, and all families assigned to treatment remained so even if they never received a text message.

Table 1
Treatment–Control Group Balance and Nonparticipating Descriptive Statistics

Variable	Panel A: Treatment (T) – Control (C) Differences			Panel B: Nonparticipants (NP) vs. Participants (P)				
	Control Mean	T – C Diff.	p-Value	Obs.	NP Mean	NP – P Diff.	p-Value	Obs.
Female	0.490	-0.011	0.685	1,137	0.494	0.013	0.400	13,911
Black	0.156	0.037	0.370	1,137	0.117	-0.061	0.000	13,911
ELL	0.019	-0.002	0.775	1,137	0.016	-0.004	0.328	13,911
IEP	0.131	0.006	0.743	1,137	0.146	0.014	0.202	13,911
Baseline math	0.000	0.046	0.538	1,137				
Baseline reading	0.000	0.014	0.842	1,137				
Suspended last year	0.201	0.006	0.660	1,137	0.198	-0.015	0.218	13,911
Baseline parent logins	15.261	-0.675	0.827	1,137	13.391	-1.960	0.183	12,943
Baseline parent ever logs in	0.359	0.022	0.537	1,137	0.314	-0.049	0.001	12,943
Baseline student logins	93.544	-2.014	0.492	1,137	91.961	-3.909	0.119	12,943
Baseline student ever logs in	0.961	-0.007	0.676	1,137	0.939	-0.027	0.000	12,943
Baseline GPA	2.819	0.010	0.784	1,137	2.844	-0.040	0.130	13,832
Baseline percent days missed	0.056	0.004	0.160	1,137	0.063	0.001	0.753	12,908
Baseline percent classes missed	0.101	-0.008	0.520	1,131	0.099	-0.008	0.043	12,859
Parents in the household	1.770	-0.029	0.374	1,137	1.820	0.068	0.000	13,911

Notes: This table shows the balance on covariates between randomized treatment and control groups. The *p*-values are for tests of equality of means across the treatment and control group via a regression of the baseline covariate on an indicator for treatment status. Standard errors clustered at the school–grade level. All regressions include strata indicators. The last four columns show the means for nonparticipants of the study, along with differences when means are compared to the experimental sample. For six of these measures in Panel B, (baseline parent logins, baseline parent ever logs in, baseline student logins, baseline percent days missed, and baseline percent classes missed) we show the differences conditional on students who attended at least one class at baseline. This is because students with high number of classes missed at baseline tend not to enroll during the following school year. No parent of a student who would not attend a class during the intervention year enrolled in the study. Baseline math and reading scores were not available for nonparticipants.

statistical comparison between nonparticipants and participants of the study. Consent into the study is not significantly correlated with most predictors of student achievement or parental involvement: baseline GPA, baseline days missed, English language learner status, IEP status, gender, an indicator for being suspended, baseline number of parent logins to the parent portal, or baseline number of student logins. However, consent is significantly and positively correlated with an indicator for the student being Black, baseline parent ever logs in, baseline student ever logs in, student's classes missed, and number of parents in the household. There is the possibility that there is self-selection into the study, such as Black parents and parents who have fewer adult caretakers in the household. However, there is mixed evidence regarding selection based on parent engagement levels and the types of students who end up participating in the study. In a separate study, which did not require parental consent, Bergman and Rogers (2017) examine how takeup of this intervention is largely determined by opt-in versus opt-out offers to receive it. The latter results in more than 95 percent takeup and effects on achievement similar to those found in this paper.

C. Random Assignment

Random assignment was at the school-by-grade level to minimize the potential for spillovers into the control group. The data were initially collapsed at the grade-by-school level, and randomization was subsequently stratified by indicators for below-median GPA and middle versus high school grades. All school employees were blinded to the randomization process.

We also randomized which parent or guardian received the text-message alerts if we had contact information available for both the mother and father of a child, or if we had multiple listed guardians. The selected parent was the parent to whom the consent letter was addressed and who the trained personnel obtained consent from by phone.

Parents in the control group received the default level of information that the schools and the teachers provided. This included report cards that are sent home after each marking period every six to nine weeks, along with parent-teacher conferences and any phone calls home from teachers. As discussed above, all parents had access to the online gradebook.

III. Data Collection

We gathered data from multiple sources: administrative data, gradebook data, survey data, and texting data. We collected initial baseline data from administrative records on student grades, courses, attendance, race and ethnicity, English language status, and contact information. We also obtained data from the gradebook system, which includes student's grades, assignments and assignments scores, class-level attendance, and parent logins into the parent portal. These baseline data were available for all students in our sample frame. During the intervention we obtained monitoring records on the text messages (Figure 2). We used these data to track messaging stop rates, whether text messages were received by phone numbers, and the total number of text messages that went out weekly.

After the 2015–2016 school year concluded we surveyed parents. The surveys took place during June and August 2016. Initially, households were sent a letter

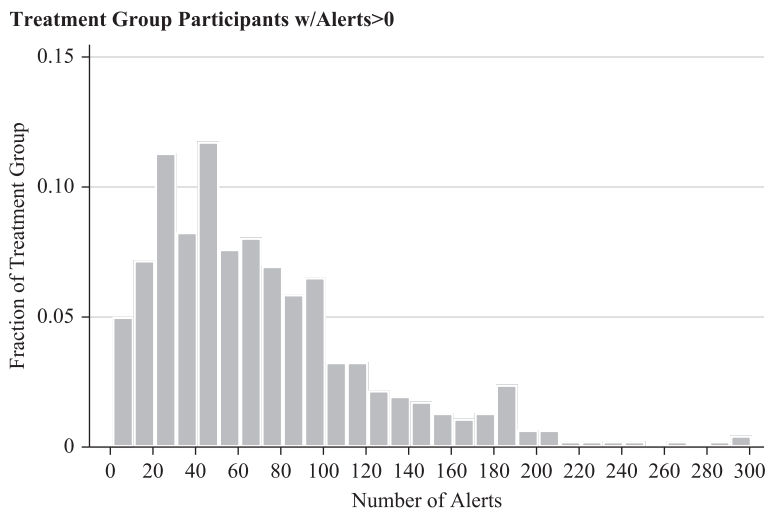


Figure 2
Distribution of Alerts for Treated Parents (Nonzero Alerts Receipt)

Notes: This figure shows the distribution of alerts for treatment parents who received at least one alert; 27.5 percent of treatment parents received zero alerts. Results derived from gradebook data.

stating that they would be called for a survey. This letter included a \$5 unconditional award as an appreciation for their participation in the study. Households were then called by trained interviewers who conducted the survey. Around this time, West Virginian residents were afflicted by severe flooding during several torrential storms in June 2016. Sadly, more than 1,000 people were left homeless in Kanawha County alone. During the summer, KCS had multiple schools declared “total losses” by the Federal Emergency Management Agency because of the flooding. As a result, we decided to mail surveys home instead of proceeding with subsequent rounds of calling. We provided a reward of \$30 for paper surveys returned postmarked by August 8, 2016. Our total response rate was 40 percent. A copy of our survey can be found in [Online Appendix C.1](#). The goal of the endline surveys was to examine parent responses to the intervention not captured by administrative data. Parents were primarily asked about their communication habits with the school in recent months, their perception of the child’s effort and achievement, and their communication and motivational habits with their child.

In the summer we obtained administrative data from the district and the gradebook application once again. These included standardized exam scores and suspension data, students’ final grades and assignment scores, daily class-level attendance, alerts received by treatment and control group, and parent and student logins into the parent portal.

A. Outcome Measures

Prior to obtaining outcome data, we registered our analysis plan, which specified sub-groups of interest, regression specifications, and primary outcomes. We specified primary outcomes to be number of classes failed, total classes attended, student retention, and math

and reading standardized test scores. Primary outcomes are derived from both the gradebook application and the KCS administrative data. Included in the gradebook data are outcomes for the number of class and daily absences and marking-period grades by course. Administrative data contained state standardized test scores in math and reading.

The standardized test scores are from the Smarter Balanced assessment, which is aligned to the Common Core. We received scaled standardized test scores for Math and ELA for 2015 and 2016 examinations. These were the first two years in which the assessment was given after the state switched from the previous standardized test in West Virginia, the WESTEST. At the time of the field experiment, students in Grades 3–11 were required to take the Smarter Balanced assessment.

We also obtained behavior-related outcomes from the gradebook application and KCS. These provided data on suspension rates, measured as the quantity of occurrences and the number of days suspended, as well as attendance measures at the class level. Following our analysis plan, we convert the latter into “number of classes present” (number of classes marked either present or tardy) because effects on retention potentially cause an increase in absences while increasing the number of days enrolled. We code suspensions into an indicator for ever being suspended.

Lastly, we use the assignment-level data to examine the effects on missed assignments, assignment scores, and class test scores. We identify tests and exams by the assignment titles containing the words “test” or “exam.” Assignment scores and test scores are standardized according to the classroom means and standard deviations for each assignment or test. We restrict the analyses to those scores three standard deviations or less away from the mean to remove outliers.¹⁷

The survey of parents was designed to examine parent and student responses to the intervention not captured by administrative and gradebook data. Parents were asked about their communication with and from the school, their perceptions about how their child was performing academically, and household behavior, such as talking with their child about their academic progress or taking privileges away as a result of their performance in school. We use a number of these survey measures, along with other gradebook and administrative measures, as secondary outcomes in this paper.¹⁸

IV. Empirical Strategy & Experimental Validity

We conduct analyses using intent-to-treat (ITT) estimates.¹⁹ We specify regressions for each parent–child pair, i , during time period t , as follows:

17. Analyses are robust to various other restrictions to handle outliers, such as excluding observations four or five standard deviations away from the mean or removing all scores from a particular assignment or exam if even one score is an outlier.

18. [Online Appendix B.4 and B.5](#) summarize all the secondary outcomes variables used in our analysis, their sources, and their construction. [Online Appendix B.6](#) summarizes the hypothesized effect on each outcome; we did not prespecify hypotheses for subgroups, however.

19. Our pre-analysis plan specified that we would run regressions to estimate treatment-on-the-treated (TOT) effects using an instrument for whether or not a participant ever received an alert. Given that control group parents could manually opt into the treatment and that some treatment group parents would never receive a text message, we believed at the time that it would be the best approach. However, during the peer review process, the feedback we received was that ITT effects are more realistic and relevant to policymakers given that

$$(1) \text{ Outcome}_{it} = \alpha_0 + \alpha_1 \text{Treatment}_i + \alpha_2 X_{it-1} + \eta_{it}$$

X_{it-1} is a set of prespecified, child-level covariates, which are fraction of days absent in the previous year, baseline GPA, an indicator for a student identified as Black, an indicator for English language learner status, an indicator for having ever been suspended in the previous year, an indicator for gender, and an indicator for having special needs. When the outcome of interest is standardized test scores, we include the baseline standardized test score as well. We impute missing covariates with the mean baseline value; we include indicators for missing to the regression as well. All regressions also include strata indicators as controls. There are 76 clusters, and we cluster standard errors at the level of treatment assignment, which is the grade level in a given school.

When looking at assignments and class test scores there are multiple observations per student—more than 70,000 assignments and 7,000 tests across the entire sample and all courses—post-treatment. The baseline control variables remain the same as above when we analyze these outcomes.

We analyze subgroups by restricting the sample to each subgroup and studying outcomes in the same way as described above. We prespecified several subgroups of interest: students with below-median GPA, students with male versus female parents or guardians, and students in middle school versus high school.

Finally, in our pre-analysis plan we hypothesized that the intervention will have positive treatment effects for our primary outcomes.²⁰ For each of these outcome measures we initially ran all tests as one-sided tests for improvements in outcomes. However, the positive, significant results we find on primary outcomes would also pass two-sided tests at conventional levels as well. As a result, for the sake of simplicity, we report all estimates using two-sided tests in this paper. Regarding secondary outcomes, we often decided to use two-sided tests initially because we did not always have a strong hypothesis about the direction of any potential effect for our secondary outcomes and subgroups.²¹ Similar to primary outcomes, we report all estimates using two-sided tests.

A. Baseline Treatment–Control Balance

Table 1, Panel A, presents baseline summary statistics for the control group, the difference in means from the treatment group, and the p -value showing the statistical significance of these differences. Demographically, the sample is 49 percent female, 16 percent Black, and the majority of students live in two-parent households. On average, students' baseline GPA is 2.8, they missed 6 percent of school days, and 20 percent

noncompliers are expected in a realistic context based on the design of the alerts. We report ITT estimates in this paper, but if interested, TOT estimates are available in the [Online Appendix](#).

20. In writing our preregistered analysis plan, we followed the Olken (2015) and Christensen and Miguel (2018) articles, which discuss preregistered analysis plans as means of transparency. Both articles state that an advantage of these analysis plans, if preregistered, is that we can specify an effect direction and one-sided hypothesis to take full advantage of statistical power. We did so in instances when we had prior evidence to suggest an effect direction. In other instances, such as many of our secondary outcomes, we were generally uncertain as to the potential direction of the effect. Regardless, our main results are statistically significant at conventional levels regardless of whether we specified one- or two-sided hypotheses.

21. [Online Appendix B.6](#) lists our secondary outcomes and hypothesized effects as stated in our pre-analysis plan.

Table 2
Measures of Attrition

	Missing Survey	Missing Suspension	Missing Math	Missing Reading
Treatment	−0.016 (0.029)	0.000 (0.010)	0.008 (0.015)	0.010 (0.015)
Percent days missed	−0.560** (0.260)	0.294 (0.200)	0.770** (0.320)	0.830** (0.330)
Baseline GPA	0.016 (0.022)	−0.003 (0.014)	0.008 (0.016)	0.015 (0.016)
Black	0.024 (0.030)	−0.007 (0.015)	−0.022 (0.014)	−0.019 (0.015)
IEP	−0.064 (0.040)	−0.008 (0.022)	0.040 (0.017)	0.042* (0.023)
Female	−0.014 (0.029)	−0.01 (0.009)	−0.011 (0.014)	−0.008 (0.015)
Baseline math score	0.034 (0.025)	0.004 (0.004)	0.026** (0.011)	0.020* (0.011)
Baseline reading score	0.012 (0.028)	0.004 (0.006)	−0.024* (0.012)	−0.021* (0.011)
Observations	1,137	1,137	1,137	1,137

Notes: This table shows the correlates of several indicators of attrition and nonresponse: survey nonresponse, missing endline GPA, and missing endline test scores. Standard clustered at the grade–school level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

were suspended during the baseline year. The contrast between by-class and full-day absences is stark: students missed 10 percent of all classes compared to 6 percent of full days. As in Bergman (2016), many more students have logged into the online gradebook portal than parents (96 percent and 36 percent, respectively).

Randomization appears to have created a treatment and control group that are similar in terms of observable variables; no treatment–controls differences are statistically significant at the 10 percent level. We also regress baseline covariates on our treatment indicator and conduct an F -test for whether these baseline covariates are jointly equal to zero. The test cannot reject that the coefficients on these covariates are jointly equal to zero (p -value = 0.61).

B. Attrition and Nonresponse

There are several sources of attrition and nonresponse in this study: missing academic outcomes, missing behavior outcomes, and survey nonresponse. A particular concern is

whether there is differential attrition by treatment status, which would invalidate our ability to make causal inferences from the data.

Table 2 shows the effect of treatment status on several measures of attrition as well as other correlates of attrition. The first column shows there is no treatment effect on the likelihood a parent responded to the survey: the point estimate is both small and statistically insignificant. Academic and demographic characteristics are generally poor predictors of survey response as well, with the exception of “percent of days missed” during the baseline academic year, which is significant at the 5 percent level. This is encouraging because it provides some suggestive evidence that our survey sample may be representative of many families in the study.

This pattern generally is also true across the remaining indicators of missing data: school suspensions, math scores, and reading scores. There are no treatment effects on any of these indicators. Only the percent of days missed the previous year is a strong predictor of missing math and reading scores, which is not surprising given that attendance the previous year predicts having measures in the current year. There are no significant predictors of missing suspension data. Overall, there is no evidence of differential attrition or nonresponse by treatment status. Additionally, attrition from course taking will be an outcome of retention analyzed below. We define retention in the district as a student taking at least one course the first semester after the intervention.

V. Results

A. Descriptive Results

We begin by describing the amount of communication between parents, children, and schools as well as parents’ beliefs about their children’s performance and their correlates. Our survey asks parents about the frequency at which they are contacted by their child’s school regarding their child’s academics. Results of the survey show that nearly 50 percent of parents heard from the school less than once every three months, about 15 percent heard once every two or three months, about 10 percent heard from the school once a month, and 25 percent of parents heard from their child’s school at least twice per month. The contrast in aggregate responses shows the variation in families who are contacted frequently and those who are not. Surprisingly little predicts this infrequency.²² Neither GPA, nor behaviors, nor demographics significantly correlate with an indicator for hearing from the school less than once every three months. This question does not, however, assess whether parents find this communication useful.

Figure 3A presents survey results describing the frequency at which parents spoke with their child about their progress in school. Fifty-five percent of parents reported talking with their child every day about their schoolwork. Roughly 75 percent of parents talked with their child two to three times per week or more. At face value, it appears their children’s schoolwork is at the top of parents’ minds. One caveat is that this communication is self-reported, which may be subject to social-desirability bias. Parents’ conversations about schoolwork is also demonstrated in Figure 3B, which shows how often parents talk to another adult in the household about their child’s school work. For this behavior, no parent reported doing so every day, but 40 percent of respondents said they

22. See [Online Appendix A.1](#), which examines the correlates of infrequent contact in Column 1.

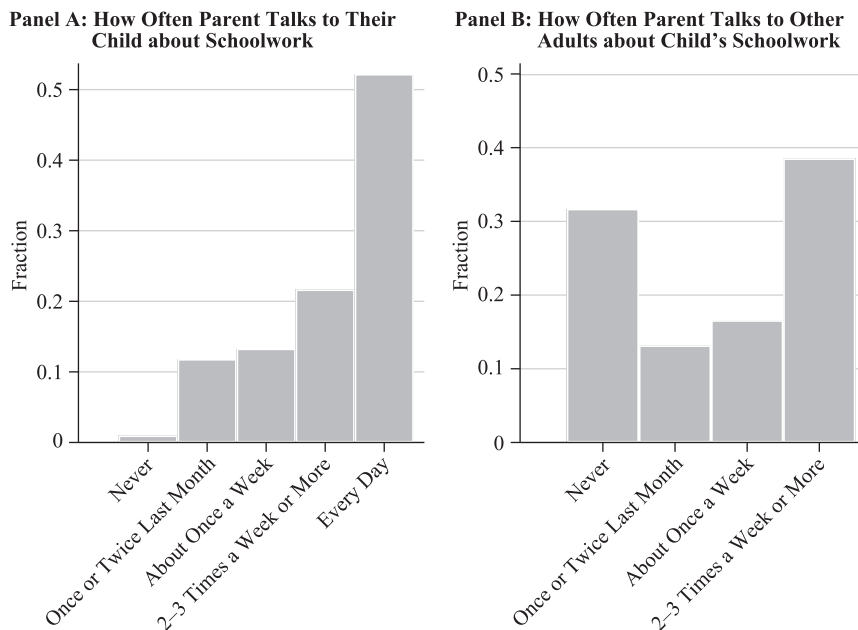


Figure 3
How Often Parent Talks to Child and Another Adult in the Household about Child's Schoolwork: Control Group Survey Results

Notes: Panel A shows the frequency of parents talking to their child about schoolwork for the control group. Panel B shows the frequency of parents talking to another adult about schoolwork for the control group. Results are from the endline parent survey.

talk with another adult two to three times per week about their child's schoolwork or grades. We also find that, unsurprisingly, two-parent households were much more likely to have intrahousehold communication about their child.²³ Little else seems to correlate with this behavior, however.

Figure 4A and 4B present control group parents' beliefs about their child's academic performance in terms of assignment completion and math grades, respectively. Figure 4A shows the number of assignments parents believed their child missed in the past semester. More than 50 percent of parents believed their child had not missed any assignments. In comparison, according to administrative data, only 20 percent of respondents' children had missed no assignments during that time frame.²⁴ In contrast, parents' perceptions about their children's math grades were more accurate: 60 percent correctly stated their child's grade in math. Around 25 percent overstated their child's grade in math. Many fewer underestimated it.

23. See [Online Appendix A.1](#) for more information.

24. [Online Appendix A.1](#) shows that inaccurate beliefs strongly and negatively correlate with their child's GPA.

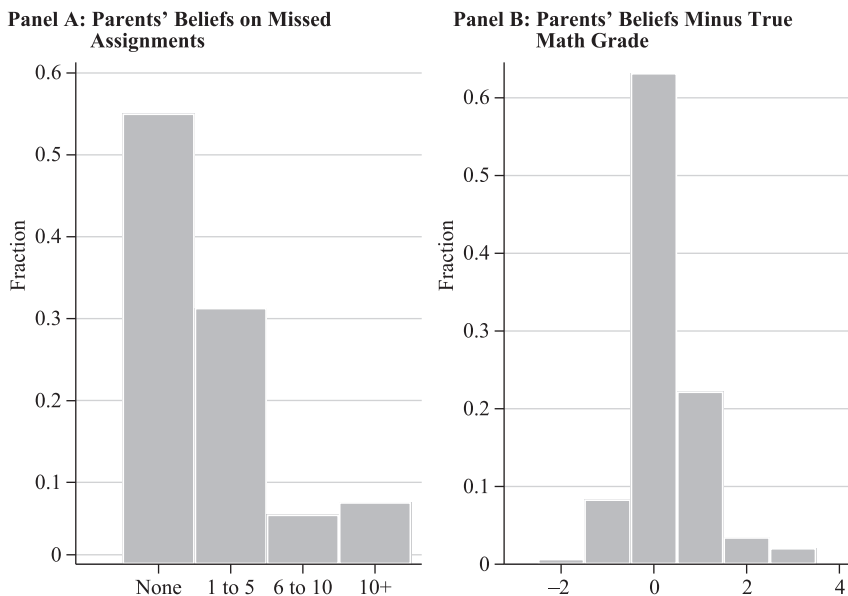


Figure 4
Control Group Parents' Beliefs about Missed Assignments

Notes: Panel A shows the fraction of parents in the control group who believe their child missed zero, one to five, six to ten, or more than ten assignments in the last semester. Panel B shows the inaccuracy of parental beliefs of math grade against actual grade. Calculations are made by subtracting actual math grade from parent's guess of student's math grade. Results to the right of zero show the fraction of parents who overestimate a student's grade, and those to the left show the fractions of parents who underestimate a student's grade. Results are calculated from the endline parent survey and gradebook data calculations.

Estimates shown in [Online Appendix A.1](#), fourth column, show a measure of the quality of communication between parents and their children: an indicator for whether parents believed it was difficult to be involved in their child's education because their child does not tell them enough about their academic progress. Forty-eight percent of parents believed their child does not disclose enough information about their academic progress for them to be easily involved in the child's education. This indicator negatively correlates with student's GPA and whether or not they are in high school. Parents with older or lower-performing children were more likely to perceive that their child is not telling them enough about their schoolwork. In results not shown, parents who reported that their children do not disclose enough also report receiving significantly fewer report cards from their child's school as well.

Overall, these descriptive statistics highlight how the information flow between parents, schools, and their children may be particularly impeded when the child is performing poorly in school. While many parents frequently spoke with their child and another adult in the household about their academic progress, nearly one-half of parents believed it would be easier to be involved in their child's education if their child told them more about their schoolwork. The latter correlates strongly with students' grades

and the receipt of report cards. In terms of parents' beliefs, parents tended to have more accurate beliefs about a student output—their grades—which is in line with what is provided on report cards. However, parents had much less accurate beliefs regarding assignment completion, which is a primary input to their child's grades. A key question is whether the automated-texting intervention studied here can increase parents' access to timely, actionable information and improve academic outcomes. The next section examines the effect of the treatment on school-to-parent communication and then looks at the effect on academic outcomes.

B. School-to-Parent Contact

We first assess whether the intervention functioned as intended. We do so by examining administrative data on alerts sent and received by parents and parent self-reported data on communication received from their child's school.²⁵ According to administrative data on alerts received, parents in the treatment group were 71 percentage points more likely to receive an alert than the control group. The distribution of alerts for nonzero alert recipients in the treatment group is also shown in Figure 2. Approximately 29 percent of treatment parents never received a single alert for missing assignments, 28 percent never received an alert for class absence, and 35 percent never received an alert for low grade.

Not every family had a cell phone to receive text messages, so compliance is imperfect. Additionally, as discussed above, all parents in the study could access the platform, and any parent could turn on our alerts by logging into the platform and turning on the alert feature. However, only two percent of the control group received any alert during the intervention period. On average, treatment group families received nearly 50 text-message alerts. Most messages were absence and assignment alerts because these were sent out weekly; families received an average of 21 of each of these alerts. Low-grade alerts went out monthly; families received about six low-grade alerts, on average.

Further, we use survey data to examine whether parents also reported receiving more contact from the school about their child's academic progress. Note that this measure includes any form of contact, such as a phone call, letter, email, or a text message from the intervention. Parents could respond: "about twice a month," "about once a month," "once every two or three months," and "less than once every three months."²⁶

In analyzing the effects of treatment assignment on these parent-reported measures of contact from the school, we find that an average of 45 percent of parents heard from their school less than three times per month about their child's progress.²⁷ Thirty-eight percent of the control group was contacted at least once per month, and treatment

25. [Online Appendix A.2](#) shows the effect of treatment status on alert receipt using administrative data. The first column shows an increase in the share of parents who received at least one alert as a result of the treatment. The second column shows the additional number of alerts that the treatment group received over the course of the school year relative to the control group. The remaining columns break the alerts down by the number of each type parents received.

26. We specified in our pre-analysis plan that we would code this into an indicator for being contacted once per month or more, but in this paper we show mutually exclusive indicators for every possible response for completeness.

27. Refer to [Online Appendix A.3](#) for estimates of parent-reported measures of contact from their children's schools.

Table 3
Effects on Primary Academic Outcomes

	Classes Failed	Classes Attended	Retained	Math Score	Reading Score
Treatment	-0.269*** (0.100)	34.190** (16.406)	0.015** (0.008)	-0.004 (0.044)	-0.056 (0.040)
Control mean	0.974	277.700	0.973	0.000	0.000
Observations	1,113	1,137	1,137	927	925

Notes: This table shows intent-to-treat (ITT) treatment effect estimates on primary academic outcomes specified in the preregistered analysis plan. All regressions include strata indicators and a set of demographic covariates described in the text. Standard errors are clustered at the grade-school level. Outcome variables are from gradebook and administrative data. Classes failed are total failed courses after treatment started. Classes attended is the numerical total of classes marked as present after treatment started. Retention is defined as taking courses one marking period after the intervention began. Math and reading scores are z scores from standardized test scores. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

increased this by 19 percentage points. There is also a 12 percentage point reduction in the likelihood that a parent reported being contacted less than once every three months.

C. Primary Academic Outcomes

In our pre-analysis plan, we specified five primary outcomes guided by the nature of the treatment, which targeted attendance, low grades, and missed assignments. These outcomes are the number of classes students failed, the number of classes attended, retention in the district, and math and reading standardized test scores.

Table 3 presents the effects on these outcomes based on ITT estimates.²⁸ Column 1 shows that control students failed a little under one course on average. Receiving text-message alerts reduced this by 28 percent or 0.27 points. Column 2 shows the effects on class attendance, which is again large and significant: students attended roughly 34 more classes than the control group, which is a 12 percent increase over the control group mean. Column 3 examines retention. About 3 percent of students from the control group did not take at least one course in the district in the semester post intervention, compared to 1.5 percent of students whose parents received alerts.

Columns 4 and 5 show that the effects on test scores are small and statistically insignificant. There are several possible reasons for this given the results discussed above. A key concern is that the exams had zero stakes for students because they have no implications for their grades or their likelihood of graduating. This issue was evident to district officials, who expressed concern that students were spending less time on the exam than expected. Smarter Balanced, the test provider, estimated that students in Grades 9, 10, and 11 need approximately 210 minutes to complete the exams at

28. Treatment-on-the-treated (TOT) estimates are shown in [Online Appendix A.4](#). As mentioned, participants assigned to treatment may not always receive the intervention, and control group users may manually opt into alerts. As a result, TOT estimates may be useful by scaling the effects of the intervention based on participants who actually received at least one alert.

each grade level. However, ninth-graders spent 80 minutes to complete the exam, tenth-graders spent 67 minutes, and 11th-graders spent 78 minutes to complete the exam, on average. The state superintendent said in response, “our students are competing with Advanced Placement exams, SAT, ACT, final exams, and then we throw in Smarter Balanced. They have told me they’re not taking it seriously.” Of particular concern was what he saw as a “lack of effort by high schools on the test.”²⁹ The state decided to discontinue using the test in future years.

We examine these claims empirically. We look at the correlations between students’ GPA and the standardized test scores by regressing math and reading state test scores on students’ baseline GPAs.³⁰ We also differentiate these correlations between middle and high school students to assess the Superintendent’s concern. The disparity in the correlations for middle school students and high school students is large: we find that the correlation between GPA and test scores is roughly one-half as large for high school students as it is for middle school students. For middle school students, the correlation between GPA and math scores and reading scores is 0.75 and 0.86, respectively. For high school students, these correlations are 0.43 for math and 0.44 for reading. This difference between middle and high schools is statistically significant (p -value < 0.001).

Second, the intervention may result in additional student effort for educational inputs that improve course grades but not learning outcomes reflected in the state test scores. The outcomes discussed above show improvements in students’ coursework and attendance. However, the curricular material covered during this additional course time may not have reflected the material covered in the exams, especially as the exams were only recently implemented in 2015. The superintendent stated that they are “working on standards-based teaching making sure all the standards are covered” (Quinn 2016c). Moreover, because the exams had only recently been introduced, no school-based accountability measures associated with the exams had been released. These reasons may attenuate the potential to impact state test scores.

D. Secondary Academic Outcomes and Behaviors and Robustness

Table 4, Panel A, presents the effects on students’ marking period course grades in more detail. Column 1 shows the effects on the number of failed courses, as before, but Columns 2–5 show the effects on the number of grades of D, C, B, and A students received as well. The intervention shifted students’ failing grades to C grades. Column 1 shows the large and significant negative effect in the number of F grades students receive, presented previously. Column 3 shows a large, positive effect—a 0.21 point increase—in the number of C grades a student receives. The coefficients on the number of B and A grades are negative and positive, respectively, but neither estimate is statistically significant. Overall, the evidence suggests that the treatment caused students to receive fewer F grades and more C grades. A primary facet of the intervention was to alert parents when

29. This made the local newspaper, the *Charleston Gazette Mail* (Quinn 2016b).

30. See [Online Appendix A.5](#) for correlations. Alternatively, we also tested for whether standardized test scores are orthogonal to other traditional demographic measures usually correlated with test scores such as race, ELL status, IEP status, gender, and number of parents in household. We find that there is very little correlation between baseline test scores and such baseline measures. In analysis not presented, a principal components analysis also formally corroborates that our test score measures are indeed orthogonal to these measures.

Table 4
Effects on Other Student-Level Academic Outcomes

	F Grades	D Grades	C Grades	B Grades	A Grades
Panel A: Student Grades					
Treatment	-0.269*** (0.100)	0.072 (0.070)	0.207** (0.084)	-0.121 (0.085)	0.180 (0.186)
Control mean	0.974	0.841	1.319	1.790	3.329
Observations	1,113	1,113	1,113	1,113	1,113
	Assignment Scores	Missed Assignments	Class Exams		
Panel B: Assignment Scores, Missed Assignments, and Class Exams					
Treatment	0.063*** (0.018)	-0.011 (0.009)	0.104*** (0.033)		
Control mean	0.018	0.119	0.008		
Observations	70,076	91,954	7,342		
	GPA	Ever Suspended	Student Logins		
Panel C: Other Student Outcomes					
Treatment	0.068* (0.045)	-0.009 (0.015)	-2.652 (11.696)		
Control mean	2.615	0.225	209.8		
Observations	1,137	967	1,137		

Notes: This table shows the ITT effect estimates on the number of each grade students received after the treatment began in Panel A, the effects on student assignment scores, the share of assignments missed, and standardized in-class test scores in Panel B, and secondary student outcomes of interest in Panel C. For Panel B, assignment and exam scores are standardized according to the control group's score for each assignment or exam. Outliers more than three standard deviations away from the mean are excluded. Missed assignments is an indicator for a missing assignment and includes assignments and exams, including assignments marked "m" for missing in the gradebook. There are multiple observations per student because there are multiple assignments or exams per student after the intervention began. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

their child's grade dips below a C, so parents know when their child is receiving a low grade. This is also consistent with the positive impacts on below-median GPA students, which we discuss when we present results on heterogeneous effects.

In Table 4, Panel B, we look closer at assignment scores, missed assignments, and class test scores. Column 1 shows that assignment scores improved by 0.06 standard

deviations over the control group. On average, the control group did not submit 12 percent of their assignments, which included classwork, exams, and homework. There is a 1.1 percent reduction in the number of missed assignments. As we show below, the effects are driven entirely by high school students, who experienced significant—both statistically and in magnitude—reductions in missed assignments.³¹

In contrast to the states' standardized test scores, scores on in-class exams increased by 0.10 standard deviations. One important difference between class tests and the standardized tests, among several, is that these scores count for students' grades and so contributed to the likelihood of a students' parent being alerted or not. The latter may have provided added incentive for students to do well on these tests as a result of the alerts.³² Comparing Column 1 with Column 3, the treatment effects are suggestively larger for exams (which are worth more points than other assignments) than assignments overall, but this difference is not statistically significant.³³

Table 4, Panel C, presents treatment effects on GPA, suspensions, and student logins. We find a positive, though insignificant, coefficient on GPA of 0.068 points. The impacts on suspensions and student logins are also insignificant. As discussed below, the effect on GPA is strong for students in high school and students with below-median GPAs at baseline. Overall, the improved assignment scores and net positive coefficient on GPA overall are encouraging. It is possible that students had held their effort constant and reallocated it toward their failing courses. The latter would not necessarily be negative given that it would result in increased credit completion, but the effects on attendance, assignment scores, and GPA provide evidence of overall net increase in student effort.

E. Mechanisms and Heterogeneity in Effects

We hypothesized that the intervention could improve the accuracy of parents' beliefs about their child's performance. To study parental beliefs, we asked parents whether they thought their child missed no assignments, one to five assignments, six to ten assignments, or more than ten assignments during the last semester. We show the effects on parents' beliefs about the number of assignments their child had missed in Table 5. Fifty-three percent of parents in the control group believed their child missed zero assignments in the last semester. The treatment reduced this belief by 11 percentage points. We can see from the remaining columns that the treatment shifted responses away from no missed assignments across the remaining categories. There is a statistically significant, six percentage point increase in the likelihood parents respond that their child has missed six to ten assignments. Only six percent of the control group believed their child missed six to ten assignments.

Figure 5A compares these beliefs about missed assignments to the number of actual missed assignments documented in the administrative data. This figure, which depicts the

31. In results not shown, all students (both treatment and control) were much less likely to miss class tests—68 percent less likely—than any other type of assignment.

32. Recent research has shown how higher stakes causes increased test scores for U.S. students (Hitt 2015; Zamorro, Hitt, and Mendez 2016; List, Gneezy, and Sadoff 2017), which can occur by reducing irrelevant answers to questions, such as writing in nonsensical words for questions with numeric answers or as using “careless” and “inconsistent” answering patterns as Hitt (2015) and Zamorro, Hitt, and Mendez (2016) find evidence of on the PISA exam.

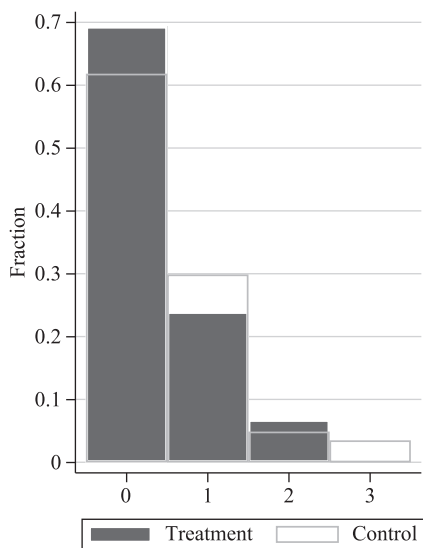
33. [Online Appendix A.6](#) shows that all of these results are robust to other approaches to outlier observations.

Table 5
Parent Beliefs about Missed Assignments

	None	1–5	6–10	>10	Don't Know
Treatment	-0.107** (0.044)	0.048 (0.036)	0.061** (0.027)	0.012 (0.025)	-0.005 (0.012)
Control mean	0.532	0.308	0.061	0.071	0.025
Observations	403	403	403	403	403

Notes: This table shows ITT estimates on parent beliefs about missed assignments. All regressions include strata indicators and a set of demographic covariates described in the text. Standard errors are clustered at the grade-school level. Outcome variables are constructed from survey results asking parents to give the number of missed assignments by students during the last semester. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A: Parents' Beliefs about Math Grade versus Truth



Panel B: Parents' Beliefs about Missed Assignments versus Truth

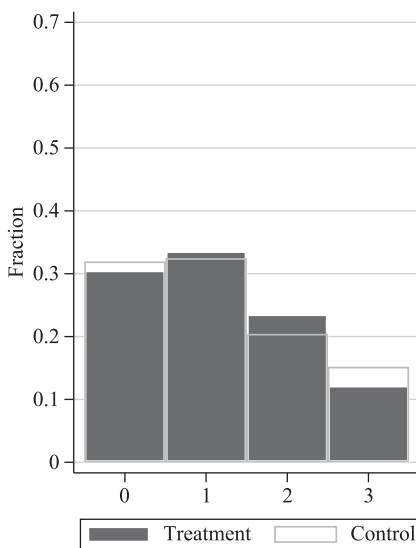


Figure 5
Treatment–Control Comparisons of Parents' Beliefs about Missed Assignments and Math Grades

Notes: Panel A shows the treatment–control comparisons of parental belief of number of missed assignments versus actual number of missed assignments. The calculations are absolute values of the inaccuracy by categorical bins in which parents estimate their child's missed assignments—zero (0), one to five (1), six to ten (2), and more than ten (3). For example, if a parent estimated that their child missed six to ten assignments, but they actually missed more than ten, they would be off by a category of one. Panel B shows the treatment–control comparisons of parental belief of their child's math grade compared to their actual grade. The calculations are absolute values of the inaccuracy by math grade GPA, based on a 4.0 scale. For example, if a child received a B, but their parent believed they received an A, the parent would be off by an absolute value of one.

Table 6
Parents' Behavioral Responses

	Contacted the School	Talked w/ Child	Parent Logins
Panel A:			
Treatment	0.119** (0.047)	0.050 (0.041)	4.988 (4.629)
Control mean	0.327	0.739	30.1
Observations	443	438	1,137
	Took Privileges	Grade Inaccuracy	Continue Texts
Panel B:			
Treatment	0.056 (0.040)	-0.133* (0.070)	0.031** (0.016)
Control mean	0.317	0.500	0.945
Observations	401	307	433

Notes: This table shows ITT estimates on parents' behavioral responses. All regressions include strata indicators and a set of demographic covariates. Standard errors clustered at the grade-school level. Outcome variables here are based on survey results and gradebook data. Panel A shows results for an indicator for whether parents contacted the school, an indicator of whether parents talked to their child about school about schoolwork or grades, and total parent logins into the parent gradebook portal. Panel B shows the results for parents taking privileges away from student in the last month of school, the difference between students' actual math grade and parents' estimated math grade, and an indicator for parents' desire to start or continue a texting service to inform them about their child's academic progress. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

absolute categorical differences in parental beliefs of missed assignments minus actual missed assignments, makes it apparent that there is no treatment effect on the accuracy of parents' beliefs about their assignment completion despite the large reduction in the share of parents who believed their child missed no assignments. Figure 5B shows a similar representation of parents' beliefs about their child's math grades relative to the grade in the administrative data. Here, there is a more visible improvement in parents' accuracy: the share of parents who accurately reported their child's grade increased by nine percentage points, and the magnitude of their errors tends to be smaller as well. We show this difference in a regression, discussed below, but a test of these distributions finds they are significantly different at the 5 percent level as well.³⁴

We also examine several additional mechanisms. First, Table 6 shows several behavioral responses to the treatment by parents as measured by survey responses. Column 1 in Panel A shows that parents were much more likely to contact the school as a result of the intervention. The share of families who contacted the school at least once

34. We use a permutation test to compare the distributions (p -value is 0.048).

over the course of the semester increased by 12 percentage points. This result is in line with Bergman (2014), who found that parents contacted the school more often in response to additional information.

The treatment effect on parent logins to the parent portal to view their child's grades is not statistically significant, as was parents speaking to their child about schoolwork daily. Parents were also asked whether they took away any of their child's privileges during the last month of school for not doing their schoolwork, which is coded as an indicator for yes or no. The effect is positive and not statistically significant, though it is close to conventional levels of marginal significance (two-sided test; p -value < 0.14). The question is worded slightly differently, but Bergman (2014) found parents were significantly more likely to take away privileges from their children.

As reported above, Column 2 of Panel B shows parents became significantly more accurate about their child's grade in math class. The third column of Panel B shows where parents would like to continue the text-message intervention. A high share—94 percent—of the control group would like to receive the intervention. The latter is not surprising, but what is encouraging is that the treatment caused a significant increase in parents' demand for the text messages of four percentage points. This suggests that any new information parents learned from receiving the alerts increased their subsequent demand for the intervention.

Finally, we also examine heterogeneity of effects on various subgroups. Given that the intervention targeted those with low grades and attendance, we are particularly interested in the subgroup of students who began the study with below-median GPAs. These heterogeneous effects also highlight some mechanisms. For instance, we see that biased beliefs about students' grades and poor parent-child communication positively correlate with students in high school and students with low GPAs (see [Online Appendix A.1](#)), and we examine how the treatment effects differ according to these subgroups that exhibit larger biases. We prespecified the following groups: students with below-median grades (by grade level), students whose father received the messages versus those whose mother received the messages, and students in middle school compared to students in high school. Tables 7 and 8 present analyses for these groups.

Table 7, Panel A, shows that students with below-median GPA failed 0.58 fewer classes, attended 42 more classes, and saw retention rates improve by three percentage points. All of these effects are significant at the 5 percent level. As before, there were no impacts on state test scores, but students' GPAs increased by 0.167 points, which is also significant at the 5 percent level. We also find that students with below-median GPAs missed 20 percent (three percentage points) fewer assignments.³⁵ As suggested above, these heterogeneous results are consistent with our survey findings showing a negative correlation between parents who believe that their child does not disclose their academic progress and their child's GPA.

Furthermore, for this subgroup of students, the treatment effect of message receipt on parents' desire to continue the intervention is 11 percentage points and significant at the 1 percent level (results not shown). This effect on the desire to continue is significantly different from the effect on parents of children with above-average GPA, who expressed no greater desire to continue the intervention than the control group (however, the mean

35. See [Online Appendix A.7](#) for estimates.

Table 7
Effects on Academic Subgroups of Interest

	Classes Failed	GPA	Classes Attended	Retained	Math Score	Reading Score
Panel A: Below Average GPA Subgroup						
Treatment	-0.578*** (0.187)	0.167** (0.070)	42.406** (18.069)	0.031** (0.013)	0.029 (0.063)	-0.003 (0.056)
Control mean	1.699	2.015	261.400	0.958	-0.414	-0.493
Observations	550	566	566	566	445	444
Panel B: High School Subgroup						
Treatment	-0.496*** (0.157)	0.171*** (0.063)	30.898* (21.519)	0.029** (0.013)	0.001 (0.060)	-0.055 (0.050)
Control mean	1.421	2.301	436.000	0.956	0.000	0.000
Observations	581	597	597	597	419	417
Panel C: Middle School Subgroup						
Treatment	-0.039 (0.110)	-0.053 (0.064)	35.613 (24.475)	0.004 (0.007)	-0.015 (0.063)	-0.059 (0.056)
Control mean	0.561	2.915	126.600	0.989	0.000	0.000
Observations	532	540	540	540	508	508

Notes: This table shows the effects by subgroups of interest, in this case students with below-median GPA at baseline in Panel A, high school students in Panel B, and middle school students in Panel C. Treatment effects are ITT estimates. All regressions include strata indicators and a set of demographic covariates described in the text. Standard errors are clustered at the grade-school level. All regressions include strata indicators. Outcome variables are from gradebook and administrative data. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

for the latter is already above 90 percent). This is further evidence of larger benefits for families whose children had lower baseline GPAs in the sample.

Panels B and C in Table 7 show that high school students were also more positively impacted than the average student. These students failed 0.5 fewer classes, attended 31 more classes, and were three percentage points more likely to remain in the district. Moreover, these effects are substantially different from the effects on middle school students, shown in Panel C. The effects for the latter group are nearly all smaller and statistically insignificant, with the exception of attendance. Additionally, we find that high school students also missed many fewer assignments (13 percent or two percentage points) than middle school students (where the coefficient is close to zero).³⁶

We examine several possible explanations for the lack of effects in middle school. First, we examine the first stage of the treatment on the alerts for the middle and high school subgroups.³⁷ The share of parents alerted in both groups is nearly identical, but

36. See [Online Appendix A.7](#).

37. See [Online Appendix A.8](#).

Table 8
Effects by Subgroups of Mother and Father Alerted

	Classes Failed	GPA	Classes Attended	Retained	Math Score	Reading Score
Panel A: Mothers Alerted Subgroup						
Treatment	-0.262** (0.130)	0.070 (0.067)	22.947* (15.944)	0.018* (0.013)	-0.011 (0.057)	0.081 (0.061)
Control mean	0.966	2.693	280.700	0.976	0.015	-0.052
Observations	423	431	431	431	345	346
Panel B: Fathers Alerted Subgroup						
Treatment	-0.356** (0.167)	0.060 (0.085)	28.294* (20.622)	0.010 (0.016)	0.072 (0.066)	-0.055 (0.072)
Control mean	0.924	2.724	272.200	0.986	0.023	0.049
Observations	319	324	324	324	266	266

Notes: This table shows the results by subgroups of interest, in this case students with either their mothers or fathers receiving the text messages. The sample is restricted to those households with two parents. Treatment effects are ITT estimates. All regressions include strata indicators and a set of demographic covariates as described in the text. Standard errors are clustered at the grade-school level. All regressions include strata indicators. Outcome variables are from gradebook and administrative data. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

parents of high school students received roughly 50 percent more alerts, which is statistically different at the 1 percent level. This difference is largely because middle school students performed much better in their grades than high school students. For instance, middle school students' GPAs were one half-standard deviation higher than high school students' GPAs at baseline (in addition to receiving fewer alerts, this also implies they had less room to improve). Second, we find that parents of high school students had more inaccurate beliefs. For instance, parents of middle school students were eight percentage points more likely to accurately recall their child's last math grades.³⁸ As discussed above, parents' beliefs about their child's math grades became significantly more accurate as a result of the intervention. In results not shown, this effect is driven almost entirely by high school students' parents and is statistically different than the effect on middle school parents' beliefs at the 10 percent level.

Table 8 shows the effects for targeting information to mothers and the fathers. While there are slight differences in effects by gender of the treated parent, the coefficients are similar in sign and show no clear pattern. Targeted fathers saw their children experience slightly better results in terms of classes failed and classes attended. In this context, we find no clear evidence that targeting one parent versus another yields different results.

Three additional groups were determined ex post to be of interest: students who were consistently absent, failed a course at baseline, and parents who never logged into the

38. Much of this difference is explained by high school students' lower GPAs.

parent portal at baseline. The first two subgroups identify lower-performing students according to outcomes we provided alerts about, and the third identifies families who may have poor access to information and inaccurate beliefs.³⁹

We find that for students with an above-median number of absences during baseline, the effects that are stronger than average.⁴⁰ These students decreased their failed courses by 0.46 courses (from a control mean of 1.44), increased class attendance by 10 percent, and increased retention by two percentage points. As before, there are still no effects on standardized test scores. For students who failed at least one course at baseline, we find evidence of larger decreases in course failures and larger increases on retention.⁴¹ However, there are insignificant, though positive, effects on attendance, math scores, and reading scores.

Given the lack of effect on standardized test scores, a key question is whether there were effects on these test scores for any subgroup. We conducted exploratory analyses to answer this question. Parents with less than a college education saw positive effects on their child's test scores, but this is a small subgroup, as the measure of parents' education is based on surveys. To corroborate this finding and expand the sample beyond the survey respondents, we linked families to census-tract-level data on college attainment and household income levels. Families living in tracts with below-median income or below-median college attainment relative to the rest of the sample saw larger effects, which were similar to the effects found for high school students, but not in terms of test scores.⁴² We do find evidence that the intervention may be particularly effective for parents who may have the least information about their children's academic performance, however. For parents who never logged into the parent portal during baseline, math scores increased by 0.08 standard deviations, though there was no significant effect on reading scores. These students' number of courses failed dropped by 0.3, GPA improved by 0.09 points, and class attended improved by 14 percent.⁴³ These results should be viewed with caution as they are exploratory. We note them here in case they prove useful in defining subgroups worthy of study in future research.

F. Second-Year Impacts and Alternative Explanations

The district permitted us to continue the intervention into a second year, but we were only able to collect data from the teacher gradebooks. By that year, KCS had

39. Parents who have never logged in to view their child's grades predicts inaccurate beliefs.

40. See [Online Appendix A.9](#).

41. See [Online Appendix A.10](#).

42. Results of income and parent educational attainment are shown in the [Online Appendix](#). [Online Appendix A.12](#) shows that students residing in census tracts where households have below-median income in Kanawha County were more positively impacted than those from census tracts with above-median income households. This is especially true for the amount of classes failed and classes attended, where students from below-median income households showed substantially larger effects. In terms of education levels, the differential effects were more modest, as seen in [Online Appendix A.13](#), which shows the differences between students living in households from census tracts with below and above-median bachelor's degree attainment in Kanawha county. The magnitude and significance of effects in classes failed and classes attended were both substantially larger for students from below-median education households. Overall, the intervention seems to affect families from low socioeconomic backgrounds more so than better-off families. However, there are still no substantial effects on test scores.

43. See [Online Appendix A.11](#).

discontinued their standardized tests in all but the 11th grade. In the summer prior to the second year of the intervention, flooding throughout the county significantly disrupted student learning. For instance, students attending one middle school and one high school only attended half days of school because flooded school sites were closed. Middle school students attended classes in the morning and high school students in the afternoon. Nonetheless, our intervention tracked students even if they moved schools within the county, though the second-year results must be interpreted within this context. The sample is reduced to 1,031 students because 12th-grade students from the first year had graduated.

We report results for the overall sample and the subgroups that appeared to drive the initial-year results: the high school sample, the below-median GPA sample, and the middle school sample. The latter saw little benefit in the first year of the intervention.

Estimates for outcomes found in the course transcripts can be found in [Online Appendix A.14](#). The outcomes include courses failed, classes attended, retention, and overall GPA during the 2016–2017 school year. The magnitude of reduction in courses failed, improvements in attendance, and retention are all similar to first-year effects and still statistically significant. The effect on overall GPA is nearly identical in magnitude and insignificant. Overall, the subgroups whose outcomes improved in the first year continued to see beneficial effects in the second year, while the middle school effects remained insignificant. Particularly encouraging is that the below-median GPA subgroup continued to see positive effects in failed courses, attendance, and retention.

Evidence regarding the ability for informational interventions to persist is mixed (see Allcott and Rogers 2014). Parents could have become habituated and unresponsive to a continual stream of information over a longer period of time. Though a second year of implementation does not provide conclusive evidence on its long-run effects, it is nonetheless encouraging that results persist into a second year. Our results cannot inform, however, any understanding of whether impacts would persist after the intervention ends.

One concern with the outcomes studied above is that teachers and students, though blinded to the random assignment, may have changed their behaviors as well. Teachers may have shifted attention or increased the frequency with which they grade assignments, and, though the intervention was randomly assigned at the school-by-grade level so that control students were always in a different grade than treated students in order to mitigate spillovers, spillovers may have occurred nonetheless.

We assess these hypotheses in several ways. First, we observe teacher logins into the gradebook, which would allow us to discern if teachers increased the frequency with which they record grades in their classrooms. Though this would be interesting if teachers had responded in this fashion, we find no impacts of the intervention on teacher logins into the gradebook for teachers who taught in treated grades compared to those teachers who did not. Second, we can observe gradebook outcomes for students who were not randomized into the intervention or the control group—the remainder of the initial recruitment sample frame. Among these students, we can compare those students who were in treated school grades (but were not treated because they were not part of the randomization process) and compare their outcomes to students who were not in treated grades (and were also not part of the randomization process). If there were evidence of spillovers, we might find that, among students who were not randomized, those in treated grades

have different outcomes than students in untreated grades.⁴⁴ We regress these gradebook outcomes on an indicator for being in a treated grade to look at these spillovers for the sample of students not randomized. We find that, across GPA, course failures, retention, and class attendance, there is no evidence of spillovers—either positive or negative.⁴⁵

Teachers also may have been wary of parents' complaints as a result of the increased school-to-parent communication and then inflated treated students' grades to avoid this reaction. We look at this possibility by examining the assignment data in greater detail. One easy way teachers could inflate grades is through "participation" grades, which may be viewed as more subjective grades relative to other assignments. Participation grades were typically marked as full credit or not—84 percent of participation grades are 100 percent scores. We therefore create an indicator for receiving a full participation grade (100 percent) or not. We find that there is no effect of the intervention on receiving a full participation score.⁴⁶ Another way teachers could have either inflated grades or reduced alerts is by marking certain assignments as "excused." We examine whether there is a change in the share of assignments marked as excused. Again, there is no significant effect, and the point estimate is essentially zero.

VI. Conclusion, Scalability, and External Validity

We helped design and implement an automated text-messaging intervention to test whether this technology can resolve parent-child information problems and improve student achievement at scale. Our intervention sends automated weekly and monthly text-message alerts to parents when their child misses a class or an assignment, or if they have a low course average. A unique feature of the intervention is that it targets class-level absences as opposed to full-day absences.

We find receiving alerts significantly reduces class-level absences, along with course failures and district attrition, though not state test scores. These effects are larger for lower-performing students and students in high school, who also miss fewer assignments and have higher in-class exam scores as a result. Notably, the effects are negligible for middle school students. Overall, students saw a shift in grades from F to C. This is particularly important given the predictive power of grades on the likelihood of dropping out (McKee and Caldarella 2016) and college success (Belfield and Crosta 2012; Vulperhorst et al. 2018; Olani 2009). Parents' beliefs about their child's grades become more accurate, but the evidence is more mixed regarding the accuracy of their child's missed assignments. We also find increases in parents' contact with schools to discuss their child's academic progress, and this is concentrated among middle school parents.

This intervention is low cost relative to other education interventions aimed at student achievement. The marginal cost of each text message is less than a fraction of one cent. Many schools already have district-wide gradebooks. Even if a school were to adopt the entire system in this study and receive training for how to use it, the cost would be \$7 per student.

44. Among students who were not part of the randomization process, we only have outcomes from the gradebook.

45. See [Online Appendix A.15](#).

46. See [Online Appendix A.16](#).

Given this low cost and policy relevance, an important question is whether this intervention would work in other contexts and whether this intervention would be adopted by parents and teachers in practice. This paper does not specifically study the adoption of the intervention by parents. However, Bergman and Rogers (2017) worked with the Washington, DC public school district to examine how varying district opt-in policies can drastically affect the adoption and, in turn, the efficacy of this particular text-message intervention. They find that when parents are randomly assigned to opt in parents by default, fewer than 5 percent of parents choose to subsequently opt out at any point during the school year. When parents are randomly assigned to actively opt in, even when this opt in is as easy as replying to a text-message offer, adoption rates are significantly lower. Bergman and Rogers also find significant reductions in courses failed and increases in student GPA, especially for high school students. The latter study did not require active consent as it was a district-initiated intervention. In the current study, less than 2 percent subsequently opted out over the course of the treatment period. Another aspect to scale is that not all districts have a unified gradebook system. For instance, Chicago Public Schools and District of Columbia Public Schools do have unified systems, but New York City does not. Our findings suggest that such a unified grade-entry system can be a valuable tool if combined with “push” information to parents, as opposed to website-only, “pull” information.

There are other open questions. For instance, we do not provide evidence on varying the frequencies, timing, and content of the information to send to parents. In our study, we target messages to low-performing students, and we do not know if tailored alerts could reach higher-performing students along other academic margins. Nor do we study how this intervention interacts with the amount of information parents already receive from various sources or how the presence of additional information may interact with the mode in which this information is sent. Though possible, we speculate that the primary concern is not whether there will be many, individually successful text-message interventions inundating parents simultaneously. Rather, an alternative scenario is that school districts begin to recognize that text messages garner parents’ attention and proceed to send even pro forma communications via text, and parents subsequently ignore all text messages. An important question for future research is how information interventions interact with the mode of communication and the existing amount of information flowing to recipients.

References

- Akresh, Richard, Damien De Walque, and Harounan Kazianga. 2016. “Evidence from a Randomized Evaluation of the Household Welfare Impacts of Conditional and Unconditional Cash Transfers Given to Mothers or Fathers.” Oklahoma State University Economics Working Paper Series. Stillwater, OK: Oklahoma State University.
- Allcott, Hunt, and Todd Rogers. 2014. “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation.” *American Economic Review* 104(10):3003–37.
- Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2017. “Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets.” *American Economic Review* 107(6):1535–63.

- Avvisati, Francesco, Marc Gurgand, Nina Guyon, and Eric Maurin. 2013. "Getting Parents Involved: A Field Experiment in Deprived Schools." *Review of Economic Studies* 81(1): 57–83.
- Belfield, Clive, and Peter Crosta. 2012. "Predicting Success in College: The Importance of Placement Tests and High School Transcripts." Community College Research Center Working Paper 42.
- Belfield, Clive, Peter Crosta, Milagros Nores, Steve Barnett, and Lawrence Schweinhart. 2006. "The High/Scope Perry Preschool Program Cost–Benefit Analysis Using Data from the Age-40 Followup." *Journal of Human Resources* 41(1):162–90.
- Bergman, Peter. 2014. "Parent–Child Information Frictions and Human Capital Investment: Evidence from a Field Experiment." Teachers College Columbia University Working Paper. New York: Columbia University.
- . 2016. "Technology Adoption in Education: Usage, Spillovers and Student Achievement." Columbia University Teachers College Working Paper. New York: Columbia University.
- Bergman, Peter, and Todd Rogers. 2017. "Is This Technology Useless? How Seemingly Irrelevant Factors Affect Adoption and Efficacy." HKS Working Paper RWP17-021.
- Berlinski, Samuel, Matias Busso, Taryn Dinkelman, and Claudia Martinez. 2016. "Reducing Parent–School Information Gaps and Improving Education Outcomes: Evidence from High Frequency Text Messaging in Chile." Unpublished.
- Bettinger, Eric, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics* 127(3):1205–42.
- Bloom, Benjamin. 1984. "The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring." *Educational Researcher* 13(6):4–16.
- Bonilla, Sheila, Sarah Kehl, Kenny Y.C. Kwong, Tricia Morphey, Rital Kachru, and Craig Jones. 2005. "School Absenteeism in Children with Asthma in a Los Angeles Inner City School." *Journal of Pediatrics* 147(6):802–6.
- Bursztn, Leonardo, and Lucas Coffman. 2012. "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas." *Journal of Political Economy* 120(3):359–97.
- Carrell, Scott, and Bruce Sacerdote. 2017. "Why Do College-Going Interventions Work?" *American Economic Journal: Applied Economics* 9(3):124–51.
- Castleman, Benjamin, and Lindsay Page. 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* 115:144–60.
- . 2016. "Freshman Year Financial Aid Nudges." *Journal of Human Resources* 51(2):389–415.
- . 2017. "Parental Influences on Postsecondary Decision Making." *Educational Evaluation and Policy Analysis* 20(10):1–17.
- Christensen, Garret, and Edward Miguel. 2018. "Transparency, Reproducibility, and the Credibility of Economics Research." *Journal of Economic Literature* 56(3):920–80.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity*. OE-38001 and Suppl. Washington, DC: U.S. Department of Health, Education, and Welfare, Office of Education.
- Cook, Philip, Kenneth Dodge, George Farkas, Fryer Roland, Jonathan Guryan, Jens Ludwig, Susan Mayer, Harold Pollack, and Laurence Steinberg. 2015. "Not Too Late: Improving Academic Outcomes for Disadvantaged Youth." Institute for Policy Research Northwestern University Working Paper WP-15-01. Evanston, IL: Northwestern University.
- Cullen, Julie Berry, Steven D. Levitt, Erin Robertson, and Sally Sadoff. 2013. "What Can Be Done To Improve Struggling High Schools?" *The Journal of Economic Perspectives* 27(2):133–52.

- Cunha, Flavio, Irma Elo, and Jennifer Culhane. 2013. "Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation." NBER Working Paper 19144. Cambridge, MA: NBER.
- Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97(2):31–47.
- Dizon-Ross, Rebecca. 2019. "Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments." *American Economic Review* 109(8):2728–65.
- Doepke, Matthias, and Michele Tertilt. 2011. *Does Female Empowerment Promote Economic Development?* Washington, DC: The World Bank.
- Doss, Christopher, Erin Fahle, Susanna Loeb, and Benjamin York. 2018. "More Than Just a Nudge: Supporting Kindergarten Parents with Differentiated and Personalized Text Messages." *Journal of Human Resources* 54(3):567–603.
- Duflo, Esther. 2012. "Women Empowerment and Economic Development." *Journal of Economic Literature* 50(4):1051–79.
- Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27(2):109–32.
- Dye, Ronald. 1985. "Disclosure of Nonproprietary Information." *Journal of Accounting Research* 23(1):123–45.
- Fryer, Roland. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." *Quarterly Journal of Economics* 126(4):1755–98.
- . 2017. "The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments." In *Handbook of Economic Field Experiments*, Volume 2, ed. Abhijit Vinayak Banerjee and Esther Dufló, 95–322. New York: Elsevier.
- Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344(6187):998–1001.
- Heckman, James. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312(5782):1900–1902.
- Heckman, James, Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94(1):114–28.
- Hitt, C.E. 2015. "Just Filling in the Bubbles: Using Careless Answer Patterns on Surveys as a Proxy Measure of Noncognitive Skills." University of Arkansas Department of Education Reform Working Paper 2015-06. Fayetteville, AR: University of Arkansas.
- Hoxby, Caroline, and Sarah Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." Stanford Institute for Economic Policy Research Discussion Paper 12-014. Stanford, CA: Stanford.
- Kinsler, Josh, and Ronni Pavan. 2016. "Parental Beliefs and Investment in Children: The Distortionary Impact of Schools." Unpublished.
- Kraft, Matthew, and Todd Rogers. 2014. "The Underutilized Potential of Teacher-to-Parent Communication: Evidence from a Field Experiment." HKS Working Paper RWP14-049.
- List, John, Uri Gneezy, and Sally Sadoff. 2017. "Measuring Success in Education: The Role of Effort on the Test Itself." NBER Working Paper 24004. Cambridge, MA: NBER.
- Mayer, Susan, Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2015. "Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together (PACT) Intervention." NBER Working Paper 21602. Cambridge, MA: NBER.
- McKee, Todd, and Paul Caldarella. 2016. "Middle School Predictors of High School Performance: A Case Study of Dropout Risk Indicators." *Education* 136(4):515–29.
- Nores, Milagros, and Steven Barnett. 2010. "Benefits of Early Childhood Interventions across the World: (Under)Investing in the Very Young." *Economics of Education Review* 29(2):271–82.

- Olani, Aboma. 2009. "Predicting First Year University Students' Academic Success." *Electronic Journal of Research in Educational Psychology* 7(3):1053–72.
- Olds, David. 2006. "The Nurse–Family Partnership: An Evidence-Based Preventive Intervention." *Infant Mental Health Journal* 27(1):5–25.
- Olken, Benjamin. 2015. "Promises and Perils of Pre-Analysis Plans." *Journal of Economic Perspectives* 29(3):61–80.
- Oreopoulos, Philip, and Uros Petronijevic. 2018. "Student Coaching: How Far Can Technology Go?" *Journal of Human Resources* 53(2):299–329.
- Page, Lindsay, Benjamin Castleman, and Katharine Meyer. 2016. "Customized Nudging to Improve FAFSA Completion and Income Verification." Unpublished.
- Quinn, Ray. 2016a. "WV Testing Panel Eyes Replacing Smarter Balanced with ACT." *Charleston Gazette Mail*, January 5. https://www.wvgazettemail.com/news/education/wv-testing-panel-eyes-replacing-smarter-balanced-with-act/article_cd6b1719-a43a-52a7-9d70-bbcbe7878e15.html (accessed October 22, 2019).
- . 2016b. "WV Scores Rise on Smarter Balanced, White State Will Continue to Use." *Charleston Gazette Mail*, August 10. https://www.wvgazettemail.com/news/education/wv-scores-rise-on-smarter-balanced-which-state-will-continue/article_ffb2575e-148a-503f-a47c-c15be0d1cc7f.html (accessed October 22, 2019).
- . 2016c. "WV School-by-School Test Results Released." *Charleston Gazette Mail*, September 28. https://www.wvgazettemail.com/news/education/wv-school-by-school-test-results-released/article_e25bcd52-1c94-5a47-ac7a-3d94ea6a797d.html (accessed October 22, 2019).
- Rogers, Todd, and Avi Feller. 2016. "Reducing Student Absences at Scale." Unpublished.
- Shin, Hyun Song. 1994. "The Burden of Proof in a Game of Persuasion." *Journal of Economic Theory* 64(1):253–64.
- Todd, Petra, and Kenneth Wolpin. 2007. "The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps." *Journal of Human Capital* 1(1):91–136.
- Vulperhorst, Jonne, Christel Lutz, Renske de Kleijn, and Jan van Tartwijk. 2018. "Disentangling the Predictive Validity of High School Grades for Academic Success in University." *Assessment & Evaluation in Higher Education* 43(3):399–414.
- Whitney, Camille, and Jing Liu. 2017. "What We're Missing: A Descriptive Analysis of Part-Day Absenteeism in Secondary School." *AERA Open* 3(2). <http://dx.doi.org/10.1177/2332858417703660>
- Yoong, Joanne, Lila Rabinovich, and Stephanie Diepeveen. 2012. "The Impact of Economic Resource Transfers to Women versus Men: A Systematic Review." Unpublished.
- York, Benjamin N., Susanna Loeb, and Christopher Doss. 2018. "One Step at a Time: The Effects of an Early Literacy Text-Messaging Program for Parents of Preschoolers." *Journal of Human Resources* 54(3):537–66.
- Zamarro, Gema, Collin Hitt, and Ildefonso Mendez. 2016. "When Students Don't Care: Reexamining International Differences in Achievement and Non-Cognitive Skills." University of Arkansas Department of Education Reform Working Paper Series 2016-18. Fayetteville, AR: University of Arkansas.