

Cortes, Kalena E.; Fricke, Hans; Loeb, Susanna; Song, David S.; York, Benjamin N.

Working Paper

When Behavioral Barriers Are Too High or Low: How Timing Matters for Parenting Interventions

IZA Discussion Papers, No. 12416

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Cortes, Kalena E.; Fricke, Hans; Loeb, Susanna; Song, David S.; York, Benjamin N. (2019) : When Behavioral Barriers Are Too High or Low: How Timing Matters for Parenting Interventions, IZA Discussion Papers, No. 12416, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/202762>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPER SERIES

IZA DP No. 12416

**When Behavioral Barriers Are Too High or
Low: How Timing Matters for Parenting
Interventions**

Kalena E. Cortes
Hans Fricke
Susanna Loeb
David S. Song
Ben York

JUNE 2019

DISCUSSION PAPER SERIES

IZA DP No. 12416

When Behavioral Barriers Are Too High or Low: How Timing Matters for Parenting Interventions

Kalena E. Cortes

Texas A&M University, IZA and NBER

Hans Fricke

Stanford University and IZA

Susanna Loeb

Brown University and NBER

David S. Song

Stanford University

Ben York

ParentPowered Public Benefit Corporation

JUNE 2019

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

When Behavioral Barriers Are Too High or Low: How Timing Matters for Parenting Interventions*

The time children spend with their parents affects their development. Parenting programs can help parents use that time more effectively. Text-messaging-based parenting curricula have proven an effective means of supporting positive parenting practices by providing easy and fun activities that reduce informational and behavioral barriers. These programs may be more effective if delivered during times when parents are particularly in need of support, such as after work, or, alternatively when parents have more time to interact with their child, such as on a day off of work. This study compares the effects of an early childhood text-messaging program sent during the weekend to the same program sent on weekdays. We find that sending the text messages on the weekend is, on average, more beneficial to children's literacy and math development. This effect is particularly strong for initially lower achieving children, while the weekday texts show some benefits for higher achieving children on higher order skills. These results are consistent with the hypothesis that the parents of lower achieving students, on average, face such high barriers during weekdays that supports are not enough to overcome these barriers, while for parents of higher achieving students, weekday texts are more effective because weekdays are more challenging, but not so difficult as to be untenable for positive parenting. In sum, the findings suggest that parenting support works best when parents have time, attention, and need.

JEL Classification: I21, I24, J18

Keywords: text messaging, parental engagement, literacy and reading skills, math skills, and parent-child activities

Corresponding author:

Kalena E. Cortes
The Bush School of Government and Public Service
Texas A&M University
4220 Tamu, 1049 Allen Building
College Station, TX 77843
USA
E-mail: kcortes@tamu.edu

* We thank Hoyt Bleakley, Susan Dynarski, Brian Jacob, Lori T. Taylor, Christina Weiland, and seminar and conference participants at the University of Michigan – The Ford School and the Association for Education Finance and Policy for helpful feedback. This research was supported by the Spencer Foundation. Erika Byun and J.B. Horsley provided outstanding research assistance. Any errors are attributable to the authors. Institutional support from Brown University, Stanford University, Stanford's Center for Education Policy Analysis (CEPA) Labs, and Texas A&M University are also gratefully acknowledged.

INTRODUCTION

Parental time investments impact children's development (Greenman, Bodovski, and Reed, 2011; Del Boca, Flinn, and Wiswall, 2014; Thomsen, 2015; Del Bono, Francesconi, Kelly, and Sacker, 2017). These time investments likely vary across time, not only year to year but also month to month, week to week, and day to day within weeks. Since most parents work outside of the home, time spent with their children is limited to non-work hours and to the days they are not working (Bianchi, 2000; Stewart, 2010; Fox, Han, Ruhm, and Waldfogel, 2013; Heiland, Price, and Wilson, 2017). Responsibilities of work, household, and finances also create cognitive demands that may affect parenting practices and may vary day to day. As a result, parenting effectiveness may differ by day, and parents may systematically benefit more from parenting supports on some days than on others. On some days, parents have greater needs for suggestions of what to do with their children because they are facing greater barriers; yet, on some days, these barriers may be so great that supports are not enough to affect parent-child interactions. In this study, we test whether the effects of a text-messaged based parenting program differ depending on the days of week it is offered, and our findings show that they do.

Parents face a number of barriers that hinder their ability to provide beneficial home learning environments for their children. Working during their children's waking hour is clearly one barrier. Information about how to create supportive environments may be another. On top of these difficulties, behavioral barriers can be strong. For instance, demands of parenting can create a heavy *cognitive load*. Parents need to make many choices each day about what questions to ask their children, what to feed them, and how to respond to their questions, requests and behaviors. These day-to-day decision alone can burden parents with a heavy cognitive load, and parents may default to doing what their parents did, what their friends do, or nothing at all, leading parents to

interact sub-optimally with their children.¹ Similarly, some of the gratifying aspects of parenting occur in the future and *delayed gratification* can cause suboptimal behavior such as underinvestment of time and effort in parenting. Parents may achieve more immediate rewards from activities such as doing the dishes or talking with friends, than from interacting in educationally positive ways with their children.² Moreover, strong cognitive demands can increase self-control problems when individuals make less mindful choices because of distractions (Shiv and Fedorikhin, 1999). Hence, parents may not engage in skill-building activities that are rewarded in the long-term on days with high barriers.³ Finally, day-to-day tasks may distract parents from more distant goals and cause parents to pay *limited attention* to beneficial parenting practices.⁴

These information, time, and behavioral constraints may be more detrimental for parents from lower socio-economic backgrounds. Wealthier parents often have jobs with more flexibility, as well as greater access to child care (Galinsky, Sakai, and Wington, 2011), while parents in poorer families often have more restrictive work hours (Williams and Boushey, 2010; Bianchi, 2011). Moreover, the strains of poverty itself limit cognitive capacity for complex tasks (Shah, Mullainathan, and Shafir, 2012; Mani, Mullainathan, Shafir, and Zhao, 2013). Parents with low incomes may use more of their cognitive capacity on daily financial challenges aggravating the behavioral barriers to parenting. In part as a result, college-educated parents spend more time with their children (Sayer, Gauthier, and Furtsenberg, 2004; Guryan, Hurst, Kearny, 2008) and are able to engage more in educational activities (Hsin and Felfe, 2014). Furthermore, the default activities

¹ See Ivengar and Lepper (2000), and Bettinger, Long, Oreopoulos, and Sanbonmatsu (2012) for more discussion on cognitive load.

² See Thaler and Sunstein (2008) and DellaVigna (2009) for more discussion on delayed gratification.

³ Mayer, Kalil, Oreopoulos, and Gallegos (2018) show that interventions that regiment reading practices according to a long-term schedule of concrete goals and rewards lead to substantial increases in time parents spend reading with their children.

⁴ See Gabraix (2017) and Karlan, McConnell, Mullainathan, and Zinman (2010) for more discussion on limited attention.

for parents who themselves were raised in highly educated families may be more helpful to children than the default behaviors of other parents, even if both groups face the same behavioral challenges.

Programs aimed at improving parenting practices may be most effective if targeted at times during which parents can engage in parenting activities and interact with their children. Given that parents have the time to spend with their children, supports for spending that time well may be more effective when parents have need for support (i.e., when barriers are present) but also have a possibility of engaging (i.e., when barriers are not insurmountable). Most parenting programs are not flexible enough to accommodate timing nuances because they provide support in a concentrated manner – usually large amounts of information at a preset time, such as during an evening class. Many of these programs have shown only limited success, perhaps because they focus on information barriers and tend not to address the behavioral barriers of parenting (Aos, Lieb, Mayfield, Miller, and Pennucci, 2004; Duncan, Ludwig, and Magnusson, 2010; Karoly, Greenwood, Everingham, Houbé, Rydell, et al., 1998). More recently, text-messaging interventions have emerged as a promising alternative or supplement due to their low cost, the widespread use of mobile phones, and their ease of scalability. These interventions have been shown to positively influence both student and parent outcomes in a wide array of educational settings.⁵ Text messaging interventions can readily target particular times and days when advice may be more useful for parents.

⁵ Such interventions have been demonstrated to positively affect school and class attendance of students (Bergman and Chan, 2017; Groot, Sander, Rogers, and Bloomenthal, 2017, Robinson, Lee, Dearing, and Rogers, 2017; Rogers and Feller, 2018), the number of course credits earned in high school (Kraft and Rogers, 2015), FASFA completion (Page, Castleman, and Meyer, 2018), chronic absenteeism and parental engagement (Smythe-Leistico and Page, 2018), and college enrollment rates (Castleman and Page, 2015, 2016). Moreover, these interventions have been particularly effective for children and parents from low-income backgrounds (Bergman, 2015; Castleman and Page, 2015, 2016; Bergman and Chan, 2017).

This study asks whether the effectiveness of text-messaging-based parenting support depends on when the support is provided. Specifically, it compares the effects of an early childhood text-messaging program sent during the weekend to those of the same program sent on weekdays. The text-messaging program was developed at Stanford University and has shown positive effects on parental engagement and children’s literacy development (Doss, Fahle, Loeb, and York, in press; York, Loeb, and Doss, in press). This intervention breaks down the complexities of parenting by providing a combination of *general information* about important literacy skills and parent-child activities (i.e., “FACT” text messages), *actionable advice* with specific examples of parent-child literacy activities (i.e., “TIP” text messages), and *encouragement/reinforcement* (i.e., “GROWTH” text messages).⁶ Starting from the hypothesis that most parents face greater distractions from positive parenting during weeknights than on weekends, if the program is more effective during the week, then the mechanism is likely through reducing barriers that are higher, while if it is more effective on the weekend, then the mechanism is likely through reducing barriers that are not so high as to be insurmountable. We expect effect heterogeneity given that some parents may face higher barriers and have less strong defaults.

We randomly assign parents of pre-kindergarten children into three groups. The first group of parents receives the original program, that is, a “FACT” message on Monday, a “TIP” message on Wednesday, and a “GROWTH” message on Friday, henceforth the *Weekday* program. The second group received the “FACT”, “TIP”, and “GROWTH” messages on Friday, Saturday, and Sunday, respectively, henceforth *Weekend* program. The *Weekend* program differs from the original program (i.e., *Weekday* program) in that it sends the text messages on different days but

⁶ The “GROWTH” text message specifically aims to provide immediate, short-term gratification to parents in the form of written encouragement. Such stimuli represent self-affirmation, which has been found to be effective in the behavioral science literature (Cohen and Sherman, 2014; Hall, Zhao, and Shafir, 2014; Sweeney and Moyer, 2015).

also in that these days are consecutive. Therefore, a third group of parents receives the “FACT”, “TIP”, and “GROWTH” message on a Tuesday, Wednesday, and Thursday, respectively, to parse out the *spacing effect* from the *weekend effect*. This program henceforth will be called **MidWeek** program.

Three prior studies have evaluated the effectiveness of the **Weekday** program. The first study, York et al. (in press), compared this program to a control group and found substantial improvements of self-reported parent engagement and child literacy assessment scores. The effects were much more pronounced for children who began the school year in the bottom half of the literacy skill distribution. The second study, Doss et al. (in press), tested whether these improvements were driven by the content of the text messages or simply by the texts serving as reminders, or nudges. To do so, the authors compared the original **Weekday** program to a **Weekday** program in which they tailored the text content to the skill level of the children. The program that targeted skill levels had a greater effect that was driven by children at the bottom and top of the literacy distribution, providing evidence that the content, not just the reminders, caused the effects. The third study, Cortes, Fricke, Loeb, and Song (2018), assessed the frequency of the text messages by comparing the original three-times-per-week **Weekday** program to a program with only one “TIP” on Wednesday and to a program that added two additional “TIP” messages on Tuesday and Thursday for a total of five texts per week. The findings demonstrated that text messaging programs can supply too little or too much information. Parents prefer three times per week to the other options and children in the bottom quarter of the pre-intervention literacy distribution benefited from receiving three texts per week compared to one but did not show further improvements when receiving five texts per week.⁷

⁷ Fricke, Kalogrides, and Loeb (2018) investigate parental opt out in these texting interventions. They find that parents who are more likely to benefit from interventions are less likely to opt out. Moreover, context and encouragement

These studies provide evidence that *content* as well as *frequency* matter for the effectiveness of the texting programs. Yet we do not know whether the *timing* of the text messages matter. That is, when do parents interact with their children, have the bandwidth to implement advice on beneficial parenting practices, and have the potential to benefit from that advice? Understanding these parenting dynamics can guide program design.

We find that sending text messages with parenting support on weekends is more beneficial to children's development, on average, than sending texts on weekdays. Moreover, we find effect heterogeneity. The benefits of the weekend program were particularly pronounced for children who started pre-Kindergarten in the lower half of the baseline skill distribution on easier sub-components, such as rapid letter naming, rote counting, and shape naming, while the weekday program was somewhat more beneficial for the initially higher achieving children. These results are consistent with the hypothesis that the parents of lower achieving students, on average, face such high barriers during weekdays that supports are not enough to overcome these barriers, while for parents of higher achieving students, weekday texts are more effective because weekdays are more challenging but still not so difficult as to impede positive parenting. In sum, the findings suggest that parenting support works best when parents have time, attention, and need.

THE INTERVENTION

The text-messaging programs are designed to support parenting practices of parents' four-year-old pre-Kindergarten children over the course of eight months. ParentPowered (parentpowered.com) developed and delivered the texts based on the original *Weekday* program

along with activities reduce opt out compared to activities alone. A high quantity of texts and more complex texts also increase opt out.

created at Stanford and first used in the San Francisco Unified School District in the 2013-14 school year (York et al., in press). The *Weekend*, the *MidWeek*, and the *Weekend* programs are identical with the exception of the days of delivery.

We use the “FACT”, “TIP”, and “GROWTH” text message approach to help parents overcome barriers that might limit parental engagement with their children.⁸ First, the “FACT” text messages intend to solve *imperfect information* by providing general information about child literacy development and best parenting practices.⁹ Second, activities in the “TIP” and “GROWTH” messages aim at alleviating the *cognitive load* of parenting by presenting concrete examples of parent-child literacy activities that are easy to implement and fun. Third, the “GROWTH” text messages aim to overcome the problem of *delayed gratification* and inconsistent time preferences. With praise and encouragement, such as “*You’re doing a good job preparing your child for kindergarten*”, the messages provide immediate gratification for skill building and school readiness activities that otherwise would be only be rewarded in the long-term. Finally, the weekly frequency of the text messages reminds parents to engage in beneficial activities with their child and solves the potential problem of *limited attention*. Examples of each text message type include:¹⁰

FACT: “*Children need to know letters to learn how to read & write. Research shows that kids with good letter knowledge become good readers.*”

TIP: “*Point out the first letter in your child's name in magazines, on signs & at the store. Have your child try. Make it a game. Who can find the most?*”

GROWTH: “*Keep pointing out letters. You're preparing your child 4K! Point out each of the letters in your child’s name. Ask: What sound does it make?*”

⁸ See a more in-depth discussion of how the programs aim to overcome these barriers in York et al. (in press).

⁹ See Avery and Kane (2004), Grodsky and Jones (2007), Hastings and Weinstein (2008), Valant and Loeb (2014), Kraft and Rogers (2015), Rogers and Feller (2016), and Fricke, Grogger, and Steinmayr (2018) for more discussion on imperfect information. This literature has produced mixed results about the effect of improved information on educational outcomes and choices.

¹⁰ More examples can be found in York et al. (in press) and in Doss et al. (in press).

We developed the content of the text messages so that they provide enjoyable and easy-to-implement activities for parent-child interactions that would help build literacy skills and more general positive parenting practices.¹¹ The text messages cover a broad range of skills, including letter recognition, rhyming, reading comprehension, vocabulary, parent-child language interactions, and parent-child reading activities.¹² We structured the texting program as a spiral curriculum, such that the program begins with simple information and advice, and becomes progressively more advanced over the eight months of intervention, while also reintroducing topics for reinforcement. Most of the activities build on existing family routines and activities to minimize costs of adopting beneficial behavior. For instance, some texts leverage bath time, commuting and travel, or meals.

We conducted this randomized control trial with the parents of four-year-old pre-kindergarten students in the Dallas Independent School District (DISD) in Texas during the 2016-17 academic year.¹³ We recruited parents through DISD's existing registration process for pre-kindergarten enrollment. As all parents in DISD must use this process to enroll in pre-kindergarten, all parents received the invitation, which was available in both English and Spanish. Parents could choose to receive text messages in either English or Spanish and had the option to withdraw from the study at any time during the school year.¹⁴ Parents who chose to participate received one of the three text messaging programs from the end of September through the beginning of June.

¹¹ See Webster-Stratton (1992), Patterson, Reid, Dishion (1992), Sanders, Markie-Dadds, Tully, and Bor (2000), Van Zeijl, Mesman, Van Ijzendoorn, Bakermans-Kranenburg, Juffer et al. (2006), and Gardner, Burton, Klimes (2006) for details.

¹² See Lonigan and Shanahan (2009) for more discussion on literacy development research; Reese, Sparks, and Leyva (2010) on parenting practice; and Abbot, Lundin, and Ong (2008) and Texas Education Agency (2015) on behavior change strategies.

¹³ We used a blocked randomization design within pre-school site and preferred texting language.

¹⁴ York et al. (in press) shows that 90 percent of economically disadvantaged families had unlimited text messaging plans.

DATA

DISD is the second-largest public-school district in the state and the 14th-largest public-school district in the United States,¹⁵ serving approximately 150,000 students from kindergarten to the twelfth grade across 224 schools. Approximately 12,000 pre-kindergarten students are enrolled in 132 preschools in the DISD district.

DISD pre-kindergarten serves students with limited English proficiency and National School Lunch Program eligibility, a diverse and low-income student population. Twenty-seven percent of the students are identified as black, 68 percent as Hispanic, and two percent as white, with the remaining three percent as Asian, American Indian, Hawaiian, multi-racial, or without a reported ethnicity. Fifty-three percent of these students are classified as limited English proficiency, and 93 percent are classified as economically disadvantaged.¹⁶

Data Sources: We use information about DISD pre-kindergarten students, their teachers, and their parents. Student data come from DISD administrative student records, which include demographic information such as age, gender, race and ethnicity, and whether the child comes from an economically disadvantaged family. Our main outcomes are literacy and mathematics scores of the Circle Assessment System (henceforth Circle), which is a one-on-one assessment and is available in English- and Spanish-language versions.

Circle evaluates children's early language and literacy skills through three primary components: rapid letter naming (a one-minute timed assessment task that evaluates the child's alphabet knowledge), rapid vocabulary (a one-minute timed assessment task that evaluates the

¹⁵ The Dallas ISD encompasses the cities of Dallas, Cockrell Hill, Seagoville, Addison, and Wilmer, and parts of Carrollton, Cedar Hill, DeSoto, Duncanville, Farmers Branch, Garland, Grand Prairie, Highland Park, Hutchins, Lancaster and Mesquite.

¹⁶ Numbers are based on official DISD enrollment statistics as of June 2018. See <https://mydata.dallasisd.org/SL/SD/ENROLLMENT/Enrollment.jsp> for more information.

child's ability to name common objects), and phonological awareness (five-minute assessment that evaluates the child's understanding of sounds, including rhyming, alliteration, syllabication, and onset-rime).¹⁷ For the math outcomes, Circle evaluates children's math skills through seven components: rote counting, shape naming, number discrimination, number naming, shape discrimination, counting, and operations.

DISD implements the Circle assessments three times per year: in the beginning of the school year (September and October 2016), mid-school year (January and February 2017), and at the end of the school year (April and May 2017). Hereafter, we refer to the beginning and the end of the school year Circle assessments as Circle-1 and Circle-3, respectively. Because Circle-1 data collection preceded the beginning of the text messaging program, we use its literacy and math measures as covariates in our regression models to increase precision and to check the balance in observed child characteristics. Circle-3 is our main outcome.

Data about teachers comes from DISD administrative records that includes gender, race and ethnicity, years of experience, days absent from work during the school year. As each child can have multiple teachers, for each student we use mean characteristics of all their teachers during the school year. Lastly, we obtain data on parents from the enrollment forms, which provide information on their preferred texting language and highest educational attainment. We use data on teacher and parent characteristics as covariates in our regression models to improve precision.

Descriptive Statistics: Table 1 presents the descriptive statistics based on two samples: the randomization sample (N=4,419) and the analytic sample (N=3,664), which includes all observations that do not have missing literacy and math outcomes in both Circle-1 and Circle-3 assessments. As shown in Table 1 (panel A), most of the children in the randomization sample are

¹⁷ In the Spanish-language version of the Circle literacy test, the assessment for phonological awareness does not include onset-rime.

Hispanic (78 percent) followed by black (18 percent), two percent are white, and one percent are Asian. Ninety-five percent of the randomization sample is economically disadvantaged, and the average age among children is 4.56 years. These demographic characteristics are similar to those of the population of DISD pre-kindergarten students, except that Hispanic children are slightly overrepresented, and black and white children are slightly underrepresented.

Table 1 also shows the raw Circle-1 literacy and math test scores, which we standardize for the analysis. On average, children name 6.03 letters in one minute of the rapid letter naming assessment, identify 10.68 objects in one minute of the rapid vocabulary assessment, and correctly answer 10.39 items in the phonological awareness assessment. For reference, the maximum score in the rapid letter naming assessment is 52 (i.e., the 26 lower-case and 26 upper-case letters in the alphabet); in the rapid vocabulary assessment the maximum score is 55 objects named with vocabulary; and in the phonological awareness assessment the maximum score is 28 points across the four subtasks (i.e., nine in rhyming, seven in alliteration, seven in syllabication, and five in onset-rime). The children in our sample had rapid letter naming and rapid vocabulary scores lower than age-appropriate standards, and phonological awareness scores higher than age-appropriate standards.¹⁸ For the math test scores, on average, children in our sample score: 0.61 points on rote counting, 3.12 points on shape naming, 1.43 points on number discrimination, 1.43 points on number naming, 4.48 points on shape discrimination, 2.70 points on counting, and 0.91 points on operations. These scores are consistently lower than age-appropriate standards.¹⁹

¹⁸ According to the CIRCLE user guide of 2016-17, the benchmarks for English-speaking children between ages four and 4.5 are 8 letters for rapid letter naming, 16 items for rapid vocabulary, and 8 points for phonological awareness; for their Spanish-speaking counterparts, the benchmarks are 6 letters, 12 items, and 5 points, respectively. See Children's Learning Institute. (2016). *CIRCLE progress monitoring: User guide 2016-17*. Houston, TX: Children's Learning Institute. Retrieved from: https://cliengage.org/user-guides/CIRCLE_Progress_Monitoring_User_Guide_91416.pdf.

¹⁹ According to the CIRCLE user guide of 2016-17, the benchmarks for both English and Spanish-speaking children between ages four and 4.5 are 2 points for rote counting, 4 points for shape naming, 2 points for number discrimination, 3 points for number naming, 5 points for shape discrimination, 4 points for counting, and 3 points for operations.

Panels B and C of Table 1 show parent and teacher characteristics, respectively. On average, parents tend to have relatively low levels of educational attainment. In the randomization sample, 32 percent of parents have less than a high school degree, 32 percent have a high school degree, and only 28 percent have completed at least some college. Fifty-seven percent of parents chose to receive text messages in English and the remaining 43 percent chose to receive them in Spanish. Three-quarters of the teachers are female, with an average teaching experience in the district of 7.79 years. The racial and ethnic demographics of teachers differ from their students: 43 percent are black, 22 percent are Hispanic, and 32 percent are non-Hispanic white.

EMPIRICAL STRATEGY

We estimate the treatment effect of the *MidWeek* and *Weekend* programs compared to the control *Weekday* program using the following model specification:

$$y_{is} = \alpha + \beta_1 \cdot \text{MidWeek}_{is} + \beta_2 \cdot \text{Weekend}_{is} + \delta \cdot X_{is} + \gamma_s + \varepsilon_{is} \quad (1)$$

where y_{is} is the outcome of interest of child i in randomization block s . The outcomes are Circle-3 literacy and mathematics outcomes. We standardize both overall scores and individual components (e.g., rapid letter naming, shape discrimination, etc.) to have a standard deviation of one and a mean of zero. We standardize the literacy assessments within assessment language because of language specific assessment differences. The variables MidWeek_{is} and Weekend_{is} are binary indicators for whether a parent received the *MidWeek* or the *Weekend* programs, respectively, in comparison to the original *Weekday* program. X_{is} is a vector of covariates that includes the child, parent, and teacher characteristics. These covariates include: the child's age, gender, race and ethnicity, whether they are economically disadvantaged, and their Circle-1 literacy and mathematics outcomes, which were measured before the text messaging programs began; the parents' highest educational attainment; and the averaged characteristics of a child's

teachers, including their gender, race and ethnicity, and days absent from work during the school year. γ_s indicates randomization block fixed effects and ε_{is} is a child-level error-term. We cluster standard errors at the randomization block level.

The coefficients of interest are β_1 and β_2 as they measure the causal effects of receiving the *MidWeek* or the *Weekend* programs compared to the original *Weekday* program. To explore heterogeneity of these treatment effects, we also estimate equation (1) in the lower half and upper half of the respective Circle-1 distribution. For literacy outcomes, we split the sample based on the distribution of overall literacy score, and for math outcomes, we split the sample based on the distribution of the overall math score.

Randomization Checks: In a randomized control trial, the characteristics of the treatment groups should not differ systematically. Systematic differences among the treatment groups could potentially bias the estimated treatment effects. Therefore, we conduct a series of randomization checks for all covariates to evaluate covariate balance. That is, we estimate the following randomization block fixed effect models for each of our child, parent, and teacher covariates:

$$X_{is} = \alpha + \beta_1 \cdot MidWeek_{is} + \beta_2 \cdot Weekend_{is} + \gamma_s + \varepsilon_{is} \quad (2).$$

Small and statistically insignificant values for the coefficients of the treatment indicators β_1 and β_2 would indicate a successful randomization. We present results of the randomization checks in Table 2 separately for the randomization sample and the analytic sample. The results provide evidence that the randomization was successful. Of the 84 point estimates using the randomization sample, only three coefficients are significant at the five percent level or less, and two coefficients are significant at the ten percent level or less. Of the 64 point estimates using the analytic sample, one coefficient is significant at the five percent level or less, and three coefficients are significant at the ten percent level or less. This number of statistically significant point

estimates can be expected to be produced by chance. Moreover, tests for joint significance using seemingly unrelated regressions, as shown in the bottom of Panel C of Table 2, fail to reject the null hypothesis that the coefficients of the treatments across the covariates are jointly equal to zero.

Attrition Analysis: We restrict our analytic sample to include only the children that have all Circle-1 and Circle-3 literacy and math components. If attrition is related to the treatment conditions, the estimated effects would be biased. For example, if assessment completion is higher for low-performing children whose parents received the *Weekend* program, compared to children whose parents received the *Weekday* program, then the estimated effect of the *Weekend* program compared to the *Weekday* program could be biased upward. To assess systematic attrition differences between the programs, we use the following a randomization block fixed effects regression model:

$$A_{is} = \alpha + \beta_1 \cdot MidWeek_{is} + \beta_2 \cdot Weekend_{is} + \delta \cdot X_{is} + \gamma_s + \varepsilon_{is} \quad (3)$$

where A_{is} is a binary variable, which equals 1 if an observation is not included in the analytic sample and 0 otherwise. Statistically insignificant values for the coefficients β_1 and β_2 would indicate that neither treatment group is more likely to experience attrition. We also estimate this model in the upper and lower halves of the Circle-1 literacy and mathematics distribution. Table 3 shows the estimates for this analysis. The results provide evidence that selective attrition is not a problem in our study. None of the estimated coefficients of the *MidWeek* and *Weekend* programs are statistically significant, nor are these coefficients significantly different from each other.

RESULTS

Table 4 presents results for aggregate Circle-3 scores. Panel A reports the treatment effects on literacy outcomes. Although we do not find statistically significant differences for the overall sample, we do find that children who start the year in the lower half of literacy distribution benefit

more from the *Weekend* program. Initially lower-achieving children whose parents received the text messages over the weekend (e.g., Friday, Saturday, and Sunday) scored, on average, 0.061 standard deviations (SDs) higher than similar children whose parents received the texts on Monday, Thursday, and Friday. Since the coefficient of the *Weekend* program is also significantly larger in magnitude than of the *MidWeek* program, this result shows that sending texts on the weekend rather than on consecutive days is more beneficial to lower-skilled children.

Panel B of Table 4 provides the results for the math assessment. In this case, the *Weekend* program is significantly more effective than the original *Weekday* program for the full sample, though the effect is again driven only by initially lower achieving children. Children whose parents received text messages on Friday, Saturday, and Sunday scored, on average, 0.032 SDs higher on the math assessment than children whose parents received texts during the week. Children who were initially in the lower half of test scores showed a differential gain of 0.053 SDs in their math score if their parents received texts on weekends, while children in the upper half of the distribution showed no such differential improvements.

To further explore the program effects, we assess the results for the literacy and math assessment components. Table 5 reports the results for the three components of Circle-3 literacy: rapid letter naming, rapid vocabulary, and phonological awareness. We find that the results for literacy are mostly driven by improvements in rapid letter naming, the least advanced skill. Overall, children whose parents received the *Weekend* program scored, on average, 0.112 SDs higher in rapid letter naming than children whose parents received the *Weekday* program. These gains in literacy are again more pronounced for children in the lower half of Circle-1 literacy scores. Lower-skilled children's rapid letter naming scores, on average, increased by 0.167 SDs, while the scores of children in the upper half only marginally increased by 0.073 SDs compared

to children in the *Weekday* program. Furthering evidence of heterogeneity, initially higher performing children whose parents received the *Weekend* program scored lower in phonological awareness, which is the more difficult skill, than children whose parents received the *Weekday* program.²⁰

Table 6 presents results for the math components: rote counting, shape naming, number discrimination, number naming, shape discrimination, counting, and operations. The results provide evidence that rote counting and shape naming drive the overall math effects reported in panel B of Table 4. Children whose parent received the *Weekend* program surpassed children whose parents received the *Weekday* program by 0.069 SDs and 0.102 SDs, respectively. The treatment effects are greater for children who started the school year in the lower half of Circle-1 overall math, as these children had math gains of 0.093 SDs in rote counting and 0.148 SDs in shape naming. The *MidWeek* program has a significant positive effect on shape naming (0.060 SDs), which is largely driven by children in the lower half of the math baseline distribution (0.124 SDs). This suggests that spacing drives some of the weekend effects on the math components. Finally, the *Weekend* and the *MidWeek* programs both appear to reduce math scores on operations for children who started the school year in the upper half of the math skill baseline distribution by 0.080 and 0.099 SDs, respectively, compared to the *Weekday* program.

Overall, these results suggest that sending text messages to parents during the weekend is more beneficial to the skill development of their children than sending text messages during the week, especially for initially lower performing students. However, our results also suggest that

²⁰ Most of the sample of parents had a high school degree or less (69 percent of the sample). If we assess the effects separately for this group relative to more educated parents, we find consistent though not as strong results. The point estimates for the differential effects of weekend tests are 0.25 SDs for low educated parents and -0.50 SDs for higher educated parents for literacy. The corresponding numbers for math are 0.02 SDs and 0.01 SDs.

weekend texts may not be as effective as weekday texts for children in upper half of the baseline skill distribution and for some higher order skills, such as, phonological awareness and operations.

ROBUSTNESS CHECKS

All results reported from Tables 4 through 6 use pre-school site fixed effects as well as covariates that capture child, parent, and average teacher characteristics. However, in a randomized control trial in which randomization has been successful, the characteristics of study participants should not significantly differ between different treatment arms. In turn, this means that we do not expect treatment effects to change when covariates are removed from the models used for the above analysis. With this in mind, we report in Table A1 some robustness checks of our main results reported in Table 4 by estimating two additional models without control variables (shown in models 2 and 3). As observed in Table A1, the point estimates remain largely stable across model specifications.

CONCLUSION

This study addresses the question of when parenting support works best in an early childhood text messaging experiment and speaks to the importance of behavioral barriers in parenting. We compare three identical programs aiming to improve parent-child interactions that are sent on different days of the week. The goal is to assess whether the messages provide better supports on more challenging days, such as weekdays, or on days that parents are more likely to be free, such as weekends. Our results suggest that sending the program on weekends increases children's literacy and math development in comparison to the program during the week, driven by the differential effects for initially lower performing students. We find some weaker evidence

that the *Weekend* program may be less effective than the weekday program for initially higher performing students.

The results provide evidence that the timing of parenting support programs matters. Parents may benefit more from supports when they have the time and other distractions are not so high that they are insurmountable but are high enough that programs aimed to reduce the barriers have leverage. While we were not able to categorize parents by their work hours and competing demands, and while students initial test performance is likely only a weak or moderate proxy for parenting practices, by separating the results based on the children's initial performance we were able to provide some evidence that delivering parenting supports at times when parents have lighter cognitive loads rather than when their needs are highest may be more beneficial especially for the most challenged parents. For parents facing lesser challenges, this approach is not necessarily better and targeting towards their more difficult times may show some benefits, but this differential is not as strong.

The importance of timing has implications for the design of parenting interventions. Although concentrated parenting programs, such as home visiting programs and parenting courses, can convey the same amount if not more information than text messaging programs, parents may not be able to remember this advice in the moment they interact with their children or may struggle to implement the advice due to behavioral constraints. Text messaging programs may help overcome these constraints by providing recurring support in small doses and more closely targeting times of parental engagement in skill building activities.

As time use patterns depend on work and non-work demands, the optimal timing of program delivery differs for different parents. Hence, our finding further highlights the benefits of differentiating programs beyond individualizing the difficulty of activities (Doss et al., in press)

and provides additional evidence that flexible light-touch programs that can differentiate to parental need and child skill level may be particularly promising avenues for program improvement.

References

- Abbott, D., Lundin, J., and Ong, F. (Eds.) (2008). *California preschool learning foundations (Vol. 1)*. Sacramento, CA: California Department of Education.
- Aos, S., Lieb, R., Mayfield, J., Miller, M., and Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Washington State Institute for Public Policy.
- Avery, C., and Kane, T.J. (2004). Student perceptions of college opportunities. The Boston COACH Program. In Caroline Hoxby (ed.), *College choices: The economics of where to go, when to go, and how to pay for it*, pp. 355-394, Chicago, IL: The University of Chicago Press.
- Bergman, P. (2015). Parent-child information frictions and human capital investment: Evidence from a field experiment. *CEStifo Working Paper Series No. 5391*. Retrieved from <https://ssrn.com/abstract=2622034>.
- Bergman, P., and Chan E.W. (2017). Leveraging parents through technology: the impact of high-frequency information on student achievement. *The Journal of Human Resources*, forthcoming.
- Bettinger, E.P., Long, B.T., Oreopoulos, P., and Sanbonmatsu, L. (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), pp. 1205-1242.
- Bianchi, S.M. (2000). Maternal employment and time with children: Dramatic change or surprising continuity? *Demography*, 37(4), pp. 401-414.
- Bianchi, S.M. (2011). Changing families, changing workplaces. the future of children. *Future Child*, 21(2), pp. 15-36.
- Castleman, B.L., and Page L.C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115, pp. 144-160.
- Castleman, B.L., and Page L.C. (2016). Freshman Year Financial Aid Nudges. *The Journal of Human Resources*, 51(2), pp. 389-415.
- Cohen, G. L., Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, 65, pp. 333-371.
- Cortes, K., Fricke, H, Loeb, S, and Song, D. (2018). Too little or too much? Actionable advice in an early-childhood text messaging experiment. NBER Working Paper No. 24827. Retrieved from <http://www.nber.org/papers/w24827>.
- Del Boca, D., Flinn C., Wiswall M. (2014). Household Choices and Child Development. *The Review of Economic Studies*, 81(1), pp. 137-185.

- Del Bono, E., Francesconi, M., Kelly, Y., Sacker, A. (2016). Early maternal time investment and early child outcomes. *The Economic Journal*, 126(596), pp. F96-F135
- DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2), pp. 315-372.
- Doss, C., Fahle, E., Loeb, S., and York, B. (in press). Supporting parenting through differentiated and personalized text messaging: Testing effects on learning during kindergarten. *The Journal of Human Resources*.
- Duncan, G.J., Ludwig, J., and Magnuson, K.A. (2010). Child development. In Phillip B. Levine and David J. Zimmerman (eds.), *Targeting investments in children: Fighting poverty when resources are limited*, pp. 27-58. Chicago, IL: The University of Chicago Press.
- Fox, L., Han, W.-J., Ruhm, C., and Waldfogel, J. (2013). Time for children: Trends in the employment patterns of parents, 1967–2009. *Demography*, 50(1), pp. 25-49.
- Fricke, H., Grogger, J., and Steinmayr, A. (2018). Exposure to academic fields and major choice. *Economics of Education Review*, 64(June), pp. 199-213
- Fricke, H., Kalogrides, D., and Loeb, S. (2018). It's too annoying: Who drops out of educational text messaging programs and why. *Economics Letters*, 173, pp. 39-43.
- Gabraix, X. (2017). Behavioral Inattention. NBER Working Paper No. 24096. Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w24096.pdf>.
- Galinsky, E., Sakai, K., and Wigton, T. (2011). Workplace flexibility: From research to action. *The Future of Children*, 21(2), pp. 141-161.
- Gardner, F., Burton, J., and Klimes, I. (2006). Randomized controlled trial of parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry*, 47(11), pp. 1123-1132.
- Greenman, E., Bodovski, K., and Reed, K. (2011). Neighborhood characteristics, parental practices and children's math achievement in elementary school. *Social Science Research*, 40(5), pp. 1434-44.
- Grodsky, E., and Jones, M.T. (2007). Real and imagined barriers to college entry: Perceptions of cost. *Social Science Research*, 36(2), pp. 745-766.
- Groot, B., Sanders, M., Rogers, T., and Bloomenthal, E. (2017). I get by with a little help from my friends: Two field experiments on social support and attendance in further education colleges in the UK. *The Behavioral Insights Team*. Retrieved from: http://38r8om2xjhh125mw24492dir.wpengine.netdna-cdn.com/wp-content/uploads/2017/04/Study-Supporter-WP_April-2017.pdf
- Guryan, J., Hurst E., and Kearney M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, 22(3), pp. 23-46.

- Hall, C. C., Zhao, J., and Shafir, E. (2014). Self-affirmation among the poor: cognitive and behavioral implications. *Psychological Science*, 25(2), pp. 619-625.
- Hastings, J.S., and Weinstein, J.M. (2008). Information, school choice, and academic achievement: Evidence from two experiments. *Quarterly Journal of Economics*, 123(4), pp. 1373-1414.
- Heiland, F., Price, J. and Wilson, R. (2014). Maternal employment and time investments in children, *Review of the Economics of the Household*, 15(1), pp. 53-67.
- Hsin, A., and Felfe, C. (2014). When does time matter? Maternal employment, children's time with parents, and child development. *Demography*, 51(5), pp. 1867-1894.
- Iyengar, S., and Lepper, M. (2000). When choice is demotivating: Can one desire too much of a good thing? *Personality & Social Psychology*, 79(6), pp. 995-1006.
- Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the Top of Mind: How Reminders Increase Saving. *Management Science*, 62(12), pp. 3393-3672.
- Karoly, L.A., Greenwood, P.W., Everingham, S.S., Houbé, J., Rydell, C.P., and Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*. Santa Monica, CA: RAND.
- Kraft, M.A., and Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review*, 47(2015), pp. 49-63.
- Lonigan, C.J., and Shanahan, T. (2009). Developing early literacy: Report of the National Early Literacy Panel. National Institute for Literacy. Retrieved from <https://eric.ed.gov/?id=ED508381>.
- Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149), pp. 976-980.
- Mayer, S. E., Kalil, A., Oreopoulos, P., and Gallegos, S. (2019). Using behavioral insights to increase parental engagement: The parents and children together intervention. *The Journal of Human Resources*, forthcoming.
- Page, L. C. and Castleman, B. and Meyer, K. (2018). Customized Nudging to Improve FAFSA Completion and Income Verification. Working Paper. Retrieved from <http://dx.doi.org/10.2139/ssrn.2854345>.
- Patterson, G.R., Reid, J.B., and Dishion, T.J. (1992). *Antisocial boys: A social interactional approach*. Eugene, OR: Castalia.
- Reese, E., Sparks, A., and Leyva, L. (2010). A review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy*, 10(1), pp. 97-117.

- Robinson, C. D., Lee, M. G., Dearing, E., and Rogers, T. (2018). Reducing student absenteeism in the early grades by targeting parental beliefs. *American Educational Research Journal*. Retrieved from <http://journals.sagepub.com/toc/aer/0/0>.
- Rogers, T., and Feller, A. (2016). Reducing student absences at scale. Working Paper. Retrieved from http://scholar.harvard.edu/files/todd_rogers/files/reducing.pdf.
- Rogers, T., and Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature: Human Behavior*, 2(2018), pp. 335-342.
- Sanders, M.R., Markie-Dadds, C., Tully, L.A., and Bor, W. (2000). The triple P-positive parenting program: A comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, 68(4), pp. 624-640.
- Sayer L.C., Gauthier A.H., and Furstenberg F.F. Jr. (2004). Educational differences in parents' time with children: Cross-national variations. *Journal of Marriage and Family*, 66(5), pp. 1152-1169.
- Shah, A., Mullainathan, S., and Shafir, E. (2012). The Consequences of Having Too Little.” *Science*, 338(6107), pp. 682-685.
- Shiv, B. and Fedorikhin, A. (1999). Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making. *Journal of Consumer Research*, 26(3), pp. 278-292.
- Smythe-Leistico, K., and Page, L. C. (2018). Connect-text: leveraging text-message communication to mitigate chronic absenteeism and improve parental engagement in the earliest years of schooling. *Journal of Education for Students Placed at Risk*, 23(1-2), pp. 139-152.
- Stewart, J. (2010). The timing of maternal work and time with children. *Industrial and Labor Relations Review*, 64(1), pp. 181-200.
- Sweeney, A. M., and Moyer, A. (2014). Self-affirmation and responses to health messages: A meta-analysis on intentions and behavior. *Health Psychology*, 34(2), pp. 149-159.
- Texas Education Agency. (2015). *Texas Prekindergarten Guidelines*. Retrieved from <https://tea.texas.gov/index2.aspx?id=2147495508>.
- Thaler, R.H., and Sunstein, C.R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
- Thomsen, M.K. 2015. Parental time investments in children: Evidence from Denmark. *Acta Sociologica*, 58(3), pp. 249-263.
- Valant, J., and Loeb, S. (2014). Information, choice, and decision-making: Field experiments with adult and student school choosers. Working Paper.

Van Zeijl, J., Mesman, J., Van Ijzendoorn, M.H., Bakermans-Kranenburg, M.J., Juffer, F., Stolk, and Alink, L.R.A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing behavior problems: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 74(6), pp. 994-1005.

Webster-Stratton, C. (1992). *The Incredible Years: Basic program manual*. Seattle, WA: The Incredible Years.

Williams, J. and Boushey, H. (2010). The three faces of work-family conflict: The poor, the professionals, and the missing middle. Working Paper. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2126314.

York, B.N., Loeb, S., and Doss, C. (in press). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. *The Journal of Human Resources*.

Table 1: Summary Statistics - Means and Standard Deviations

	Randomization Sample	Circle-3 Test Sample*
Panel A: Children Characteristics		
Age (in years)	4.56 (0.29)	4.57 (0.29)
Female	0.50	0.51
Race and ethnicity		
Black	0.18	0.16
Hispanic	0.78	0.80
Asian	0.01	0.01
White	0.02	0.02
Other**	0.01	0.01
Low socioeconomic status	0.95	0.96
Circle-1 Literacy assessment scores ^a		
Rapid Letter Naming	6.03 (8.49)	6.08 (8.74)
Missing Rapid Letter Naming	0.06	0.00
Rapid Vocabulary	10.68 (6.97)	10.63 (7.22)
Missing Rapid Vocabulary	0.06	0.00
Phonetic Awareness	10.39 (5.54)	10.47 (5.65)
Missing Phonetic Awareness	0.07	0.00
Circle-1 Math assessment (raw) scores ^a		
Rote Counting	0.61 (0.64)	0.61 (0.66)
Missing Rote Counting	0.07	0.00
Shape Naming	3.12 (1.56)	3.15 (1.61)
Missing Shape Naming	0.06	0.00
Number Discrimination	1.43 (0.73)	1.44 (0.75)
Missing Number Discrimination	0.06	0.00
Number Naming	1.43 (0.73)	1.44 (0.75)
Missing Number Naming	0.06	0.00
Shape Discrimination	4.48 (1.58)	4.51 (1.63)
Missing Shape Discrimination	0.07	0.00
Counting	2.70 (1.55)	2.73 (1.61)
Missing Counting	0.07	0.00
Operations	0.91 (1.03)	0.91 (1.08)
Missing Operations	0.08	0.00
Observations	4,419	3,664

Notes: Numbers in parentheses are standard deviations for continuous variables. *Circle-3 test sample is conditioned on having Circle-1 test scores. **Includes American Indian, Native American, Pacific Islander, two or more races, or other race. ^aMissing data values imputed to be the mean.

^bMissing data values set to zero.

Table 1 (continued): Summary Statistics - Means and Standard Deviations

	Randomization Sample	Circle-3 Test Sample*
Panel B: Parental Characteristics		
Highest education level ^b		
Less than high school	0.32	0.33
High school	0.32	0.32
Some college	0.19	0.19
Associate degree	0.04	0.04
Bachelor's degree	0.03	0.03
Graduate degree	0.02	0.02
Missing education	0.07	0.08
Texting language		
English	0.57	0.55
Spanish	0.43	0.45
Panel C: Teacher Characteristics		
Average female	0.75	0.75
Average race and ethnicity		
Black	0.43	0.45
Hispanic	0.22	0.22
White	0.32	0.32
Average missing teacher demographics	0.00	0.00
Average experience (in years)	7.79	7.85
	(5.92)	(6.01)
Average days absent	9.39	9.13
	(9.16)	(9.03)
Average missing absence information	0.08	0.07
	(0.19)	(0.17)
Observations	4,419	3,664

Notes: Numbers in parentheses are standard deviations for continuous variables. *Circle-3 test sample is conditioned on having Circle-1 test scores. **Includes American Indian, Native American, Pacific Islander, two or more races, or other race. ^aMissing data values imputed to be the mean. ^bMissing data values set to zero.

Table 2: Randomization Checks - The Effect of Treatment Status on Pre-Treatment Covariates by the Randomization and Circle-3 Test Sample

Panel A: Regressions of Child Characteristics on Treatment Status (Weekday program is the omitted category in all regressions)

	Randomization Sample (N=4,419)			Circle-3 Test Sample* (N=3,664)		
	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)
Age (in years)	0.013 (0.012)	0.008 (0.011)	0.642	0.008 (0.013)	0.001 (0.012)	0.568
Female	-0.003 (0.019)	-0.012 (0.020)	0.661	-0.000 (0.021)	-0.001 (0.022)	0.971
Black	0.017 (0.010)	0.009 (0.011)	0.400	0.021* (0.011)	0.014 (0.011)	0.505
Hispanic	-0.015 (0.012)	-0.002 (0.012)	0.223	-0.019 (0.013)	-0.009 (0.013)	0.373
Asian	-0.001 (0.005)	-0.005 (0.004)	0.286	-0.004 (0.004)	-0.005 (0.004)	0.572
White	-0.006 (0.005)	-0.005 (0.005)	0.858	-0.006 (0.006)	-0.006 (0.006)	0.918
Other**	0.001 (0.001)	0.000 (0.001)	0.510	---	---	---
Low socioeconomic status	0.002 (0.008)	0.013 (0.009)	0.243	0.003 (0.007)	0.002 (0.009)	0.976
Rapid Letter Naming ^{a,b}	0.348 (0.297)	0.021 (0.313)	0.288	0.404 (0.336)	0.108 (0.385)	0.405
Rapid Letter Naming missing ^g	-0.003 (0.008)	-0.008 (0.008)	0.542	---	---	---
Rapid Vocabulary ^{a,b}	0.225 (0.265)	0.056 (0.249)	0.524	0.143 (0.295)	-0.008 (0.292)	0.602
Rapid Vocabulary missing ^g	-0.004 (0.009)	0.001 (0.009)	0.543	---	---	---
Phonological Awareness ^{a,b}	0.037 (0.199)	0.084 (0.194)	0.811	-0.050 (0.218)	0.006 (0.220)	0.795
Phonological Awareness missing ^g	-0.001 (0.009)	0.003 (0.009)	0.675	---	---	---
Rote Counting ^{a,b}	0.002 (0.022)	-0.016 (0.022)	0.407	0.001 (0.026)	-0.006 (0.026)	0.805
Rote Counting missing ^g	-0.003 (0.008)	-0.006 (0.009)	0.734	---	---	---
Shape Naming ^{a,b}	0.117** (0.059)	0.042 (0.058)	0.192	0.122* (0.068)	0.073 (0.072)	0.463
Shape Naming missing ^g	-0.005 (0.008)	-0.001 (0.009)	0.637	---	---	---
Number Discrimination ^{a,b}	0.059** (0.028)	0.025 (0.031)	0.227	0.043 (0.033)	0.009 (0.034)	0.305
Number Discrimination missing ^g	-0.006 (0.008)	-0.000 (0.009)	0.546	---	---	---
Number Naming ^{a,b}	0.098* (0.052)	0.084 (0.054)	0.792	0.117* (0.060)	0.089 (0.060)	0.643
Number Naming missing ^g	-0.001 (0.008)	-0.003 (0.009)	0.775	---	---	---
Shape Discrimination ^{a,b}	0.109** (0.054)	0.042 (0.063)	0.279	0.077 (0.063)	0.042 (0.076)	0.629
Shape Discrimination missing ^g	-0.005 (0.008)	0.003 (0.009)	0.433	---	---	---
Counting ^{a,b}	0.033 (0.054)	0.026 (0.054)	0.909	0.033 (0.063)	0.041 (0.064)	0.911
Counting missing ^g	-0.006 (0.009)	0.004 (0.010)	0.294	---	---	---
Operations ^{a,b}	0.020 (0.037)	-0.010 (0.039)	0.394	0.031 (0.042)	0.001 (0.046)	0.471
Operations missing ^g	-0.005 (0.009)	-0.000 (0.010)	0.605	---	---	---

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include pre-school site fixed effects. Standard errors are clustered at the randomization block site level. The omitted reference group in all regressions is the original texting program of Monday/Wednesday/Friday. *Circle-3 test sample is conditioned on having Circle-1 and Circle-3 literacy and math test scores. **Includes American Indian, Native American, Pacific Islander, two or more races, or other race. ^aMissing data values imputed to be the mean and all regressions include a dummy variable for categorical variables with missing values. ^bAll Circle-1 test score variables are in standard deviation units. ^cNo students included with missing information in Circle-3 test sample. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 2 (continued): Randomization Checks - The Effect of Treatment Status on Pre-Treatment Covariates by the Randomization and Circle-3 Test Sample

Panel B: Regressions of Parental Characteristics on Treatment Status (Weekday program is the omitted category in all regressions)

	Randomization Sample (N=4,419)			Circle-3 Test Sample* (N=3,664)		
	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)
Less than high school	0.007 (0.016)	0.021 (0.017)	0.356	0.003 (0.018)	0.024 (0.018)	0.263
High school	-0.007 (0.018)	-0.006 (0.017)	0.947	0.001 (0.020)	-0.004 (0.019)	0.799
Some college	0.006 (0.015)	0.005 (0.015)	0.898	0.000 (0.017)	0.013 (0.017)	0.374
Associate degree	0.003 (0.007)	0.002 (0.007)	0.822	0.002 (0.008)	-0.001 (0.008)	0.706
Bachelor's degree	-0.007 (0.007)	-0.010 (0.006)	0.590	-0.004 (0.007)	-0.013** (0.007)	0.177
Graduate degree	-0.004 (0.006)	-0.007 (0.006)	0.647	-0.000 (0.007)	-0.008 (0.006)	0.148
Missing education	0.002 (0.010)	-0.004 (0.010)	0.562	-0.002 (0.012)	-0.011 (0.011)	0.482

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include randomization block fixed effects. Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program of Monday/Wednesday/Friday. *Circle-3 test sample is conditioned on having Circle-1 and Circle-3 literacy and math test scores. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 2 (continued): Randomization Checks - The Effect of Treatment Status on Pre-Treatment Covariates by the Randomization and Circle-3 Test Samples

Panel C: Regressions of Teacher Characteristics on Treatment Status (Weekday program is the omitted category in all regressions)

	Randomization Sample (N=4,419)			Circle-3 Test Sample* (N=3,664)		
	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)
Ave. female	-0.001 (0.005)	0.000 (0.005)	0.894	-0.001 (0.006)	0.005 (0.006)	0.370
Ave. black	-0.004 (0.006)	-0.007 (0.006)	0.650	-0.003 (0.006)	-0.004 (0.008)	0.916
Ave Hispanic	-0.004 (0.008)	0.003 (0.009)	0.351	-0.006 (0.009)	-0.006 (0.010)	0.931
Ave. white	0.008 (0.006)	0.007 (0.008)	0.822	0.009 (0.007)	0.012 (0.007)	0.635
Ave. experience	0.121 (0.109)	0.068 (0.130)	0.661	0.182 (0.120)	0.096 (0.143)	0.505
Ave. days absent	-0.068 (0.198)	-0.193 (0.288)	0.623	-0.186 (0.212)	-0.362 (0.316)	0.479
Ave. missing absence info.	-0.008* (0.005)	-0.003 (0.005)	0.320	-0.002 (0.005)	-0.003 (0.005)	0.754
SUR Joint t-test (p-value) ^a	0.918	0.914		0.810	0.740	

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include pre-school site fixed effects. Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program of Monday/Wednesday/Friday.

*Circle-3 test sample is conditioned on having Circle-1 and Circle-3 literacy and math test scores. ^aSeemingly Unrelated Regression t-test of joint significance of treatments on student, parent, and teacher characteristics. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 3: Treatment Effects on Attrition in Circle-3 Test Sample -
Whether any Circle-3 Test Outcomes (Literacy and Math) are Missing

	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	Adj. R ²	N
Entire sample	-0.009 (0.010)	-0.003 (0.010)	0.519	0.56	4,419
By Pre-Treatment Literacy Score: ^a					
Lower half	0.002 (0.016)	0.001 (0.015)	0.929	0.39	2,051
Upper half	-0.011 (0.012)	-0.010 (0.013)	0.978	0.40	1,991
By Pre-Treatment Mathematics Score: ^a					
Lower half	0.002 (0.014)	-0.000 (0.014)	0.872	0.10	2,170
Upper half	-0.012 (0.013)	-0.004 (0.014)	0.489	0.09	1,818

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, days absent, and days missing) and randomization block fixed effects. Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday. ^aRegressions in samples split at medians are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 4: Treatment Effects on Overall Circle-3 Skills Assessment Test (standardized)

Panel A: Language and Literacy Skills Assessment Test Results					
	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	Adj. R ²	N
Entire sample	-0.006 (0.025)	0.026 (0.022)	0.220	0.52	3,664
By Pre-Treatment Literacy Score: ^a					
Lower half	-0.037 (0.039)	0.061* (0.034)	0.024	0.44	1,828
Upper half	0.041 (0.033)	-0.006 (0.031)	0.116	0.51	1,836
Panel B: Mathematics Skills Assessment Test Results					
	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	Adj. R ²	N
Entire sample	0.001 (0.016)	0.032** (0.015)	0.071	0.47	3,664
By Pre-Treatment Mathematics Score: ^a					
Lower half	0.007 (0.028)	0.053** (0.027)	0.144	0.41	1,970
Upper half	-0.020 (0.016)	-0.008 (0.015)	0.440	0.39	1,694

Notes: Each row represents a separate regression model (only the coefficients of the treatments status are reported). All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, days absent, and days missing) and randomization block fixed effects. Standard errors are clustered at the randomization level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday.
^aRegressions in samples split at medians are based on student's Circle-1 tests prior to the intervention.
 Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 5: Treatment Effects on Circle-3 Language
and Literacy Skills Assessment Test by Specific Sub-tests (standardized)

Panel A: Rapid Letter Naming					
	MidWeek	Weekend	Weekend vs. MidWeek	Adj. R ²	N
	Program	Program	t-test (p-value)		
Entire sample	0.023 (0.032)	0.112*** (0.029)	0.009	0.44	3,664
Lower half	-0.009 (0.050)	0.167*** (0.047)	0.002	0.37	1,828
Upper half	0.068 (0.044)	0.073* (0.043)	0.893	0.45	1,836
Panel B: Rapid Vocabulary Naming					
	MidWeek	Weekend	Weekend vs. MidWeek	Adj. R ²	N
	Program	Program	t-test (p-value)		
Entire sample	0.003 (0.034)	0.023 (0.032)	0.572	0.44	3,664
Lower half	-0.046 (0.055)	0.072 (0.047)	0.040	0.38	1,828
Upper half	0.049 (0.047)	-0.023 (0.048)	0.114	0.45	1,836
Panel C: Phonological Awareness					
	MidWeek	Weekend	Weekend vs. MidWeek	Adj. R ²	N
	Program	Program	t-test (p-value)		
Entire sample	-0.044 (0.032)	-0.057* (0.030)	0.700	0.41	3,664
Lower half	-0.055 (0.049)	-0.055 (0.049)	0.993	0.34	1,828
Upper half	0.006 (0.044)	-0.069* (0.041)	0.086	0.37	1,836

Notes: All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, days absent and days missing) and randomization block fixed effects. Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday. Regressions in samples split at medians are based on student's Circle-1 literacy tests prior to the intervention. Phonological awareness is a composite score of the following assessments: rhyming, alliteration, and syllabication. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table 6: Treatment Effects on Circle-3 Mathematics Skills Assessment Test by Specific Sub-tests (standardized)

	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	Adj. R ²	N	MidWeek Program	Weekend Program	Weekend vs. MidWeek t-test (p-value)	Adj. R ²	N
Panel A: Rote Counting						Panel E: Shape Discrimination				
Entire sample	0.030 (0.032)	0.069** (0.033)	0.212	0.29	3,664	-0.002 (0.040)	0.037 (0.037)	0.323	0.17	3,664
Lower half	0.054 (0.053)	0.093* (0.052)	0.463	0.27	1,970	-0.032 (0.068)	0.025 (0.062)	0.437	0.17	1,970
Upper half	0.010 (0.038)	0.026 (0.038)	0.643	0.18	1,694	0.013 (0.049)	0.031 (0.047)	0.662	0.14	1,694
Panel B: Shape Naming						Panel F: Counting				
Entire sample	0.060* (0.033)	0.102*** (0.034)	0.198	0.31	3,664	-0.005 (0.032)	0.042 (0.033)	0.174	0.29	3,664
Lower half	0.124** (0.060)	0.148** (0.060)	0.679	0.30	1,970	0.003 (0.057)	0.087 (0.057)	0.146	0.29	1,970
Upper half	-0.033 (0.031)	0.011 (0.031)	0.190	0.21	1,694	-0.018 (0.039)	-0.007 (0.040)	0.778	0.17	1,694
Panel C: Number Discrimination						Panel G: Operations				
Entire sample	-0.026 (0.038)	-0.006 (0.041)	0.600	0.09	3,664	-0.062* (0.033)	-0.037 (0.035)	0.478	0.30	3,664
Lower half	-0.033 (0.070)	-0.001 (0.076)	0.657	0.08	1,970	-0.050 (0.058)	-0.020 (0.055)	0.634	0.25	1,970
Upper half	-0.038 (0.037)	-0.012 (0.029)	0.441	0.14	1,694	-0.099** (0.044)	-0.080* (0.046)	0.690	0.28	1,694
Panel D: Number Naming										
Entire sample	0.006 (0.030)	0.047 (0.029)	0.177	0.39	3,664					
Lower half	-0.004 (0.051)	0.083 (0.050)	0.098	0.31	1,970					
Upper half	-0.019 (0.039)	-0.027 (0.037)	0.808	0.30	1,694					

Notes: All regressions include controls for student characteristics (age, gender, race/ethnicity, low-SES status, and Circle-1 test scores), parental characteristics (highest educational attainment), teacher characteristics (gender, race/ethnicity, years of experience, days absent and days missing) and randomization block fixed effects. Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday. Regressions in samples split at medians are based on student's Circle-1 tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table A1: Robustness Checks for Table 4 Results - Comparison of Regression Models
 Comparison of Regression Models for Treatment Effects on Overall Circle-3 Test Results (standardized)

Model 1: Table 4 Results: Randomization block fixed effects with full set of control variables including initial Circle-1 assessment scores.

Model 2: Randomization block fixed effects with full set of control variables excluding initial Circle-1 assessment scores.

Model 3: Only randomization block fixed effects and no additional control variables.

Panel A: Entire sample - Overall Literacy			
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	-0.006 (0.025)	0.016 (0.027)	0.024 (0.028)
Weekend Program	0.026 (0.022)	0.038 (0.029)	0.038 (0.030)
Weekend vs. MidWeek t-test (p-value)	0.220	0.461	0.625
N	3,664	3,664	3,664
Adj. R ²	0.52	0.31	0.25
Panel B: Entire sample - Overall Mathematics			
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	0.001 (0.016)	0.020 (0.019)	0.026 (0.020)
Weekend Program	0.032** (0.015)	0.045** (0.021)	0.046** (0.021)
Weekend vs. MidWeek t-test (p-value)	0.071	0.216	0.315
N	3,664	3,664	3,664
Adj. R ²	0.47	0.21	0.17

Notes: Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.

Table A1 (continuation): Robustness Checks for Table 4 Results -
 Comparison of Regression Models for Treatment Effects on Overall Circle-3 Test Results (standardized)

Model 1: Table 4 Results: Randomization block fixed effects with full set of control variables including initial Circle-1 assessment scores.

Model 2: Randomization block fixed effects with full set of control variables excluding initial Circle-1 assessment scores.

Model 3: Only randomization block fixed effects and no additional control variables.

Panel C: Literacy - By Pre-Treatment Literacy Score			
	Lower half		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	-0.037 (0.039)	-0.017 (0.041)	-0.003 (0.042)
Weekend Program	0.061* (0.034)	0.084** (0.039)	0.096** (0.038)
Weekend vs. MidWeek t-test (p-value)	0.024	0.026	0.032
N	1,828	1,828	1,828
Adj. R ²	0.44	0.31	0.29
	Upper half		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	0.041 (0.033)	0.078** (0.036)	0.078** (0.038)
Weekend Program	-0.006 (0.031)	-0.003 (0.037)	-0.006 (0.037)
Weekend vs. MidWeek t-test (p-value)	0.116	0.019	0.014
N	1,836	1,836	1,836
Adj. R ²	0.51	0.35	0.33
Panel D: Mathematics - By Pre-Treatment Mathematics Score			
	Lower half		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	0.007 (0.028)	0.014 (0.031)	0.022 (0.032)
Weekend Program	0.053** (0.027)	0.053* (0.031)	0.058* (0.031)
Weekend vs. MidWeek t-test (p-value)	0.144	0.235	0.281
N	1,970	1,970	1,970
Adj. R ²	0.41	0.24	0.22
	Upper half		
	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
MidWeek Program	-0.020 (0.016)	-0.018 (0.016)	-0.016 (0.017)
Weekend Program	-0.008 (0.015)	-0.007 (0.017)	-0.007 (0.017)
Weekend vs. MidWeek t-test (p-value)	0.440	0.530	0.613
N	1,694	1,694	1,694
Adj. R ²	0.39	0.27	0.26

Notes: Standard errors are clustered at the randomization block level. The omitted reference group in all regressions is the original texting program Monday/Wednesday/Friday. ^aRegressions in samples split at medians are based on student's Circle-1 literacy (or math) tests prior to the intervention. Statistical significance levels: *p<0.10; **p<0.05; ***p<0.01.