August 2023

Summer School As a Learning Loss Recovery Strategy After COVID-19:

Evidence from Summer 2022

Ian Callen
Maria V. Carbonari
Michael DeArmond
Daniel Dewey
Elise Dizon-Ross
Dan Goldhaber
Jazmin Isaacs
Thomas J. Kane
Megan Kuhfield
Anna McDonald
Andrew McEachin
Emily Morton
Atsuko Muroga
Douglas O. Staiger

CALDER
National Center for Analysis of Longitudinal Data in Education Research

Center for Education Policy Research
HARVARD UNIVERSITY

nwea

SUGGESTED CITATION
Callen, I., Carbonari, M. V., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Isaacs, J., Kane, T. J., Kuhfeld, M., McDonald, A., McEachin, A., Morton, E., Muroga, A., & Staiger, D. O. (2023). Summer School As a Learning Loss Recovery Strategy After COVID-19: Evidence from Summer 2022. Research Report. Cambridge, MA: Center for Education Policy Research, Harvard University/CALDER Working Paper No. 2901- 0823.

This research was supported by the Carnegie Corporation of New York, the Walton Family Foundation, Kenneth C. Griffin, and an anonymous foundation. CALDER is funded by a consortium of foundations (for more information about CALDER funders, see www.caldercenter.org/about-calder). We could not have drafted this report without the district leaders who generously gave their time and attention to us during a challenging school year. We are also grateful to Erin Fahle and Lisa Sanbonmatsu for assistance with various aspects of this work.

Table of Contents

ABSTRACT

To make up for pandemic-related learning losses, many U.S. public school districts have increased enrollment in their summer school programs. We assess summer school as a strategy for COVID-19 learning recovery by tracking the academic progress of students who attended summer school in 2022 across eight districts serving 400,000 students. Based on students' spring to fall progress, we find a positive impact for summer school on math test achievement (0.03 standard deviation, *SD*), but not on reading tests. These effects are predominantly driven by students in upper elementary grades. To put the results into perspective, if we assume that these districts have losses similar to those present at the end of the 2022–23 school year (i.e., approximately -0.2 *SD*), we estimate summer programming closed approximately 2% to 3% of the districts' total learning losses in math, but none in reading.

1. INTRODUCTION

Following the pandemic, many students remain behind academically, and the pace of recovery is slow (Lewis & Kuhfeld, 2023). School districts have responded with a range of academic interventions, from extended school calendars to high-dosage tutoring (Carbonari et al., 2022; Diliberti & Schwartz, 2022). Although discussions about COVID-19 recovery often focus on the promise of high-dosage tutoring (Nickow et al., 2020), summer school has also been a popular recovery strategy. "Summer learning" was the most common example of academic recovery spending in a recent analysis of 5,000 district spending plans for the American Rescue Plan Elementary and Secondary School Emergency Relief (ESSER) funds (Dimarco & Jordan, 2022b). This same analysis estimates that districts' total ESSER spending on summer programs will reach $5.8 billion by September 2024 (DiMarco & Jordan, 2022a). In a national survey from 2022, 70% of districts reported providing new or expanded summer programming because of the pandemic (Diliberti & Schwartz, 2022). The importance of summer learning has also been touted by district leaders: Alberto Carvalho, the superintendent of Los Angeles Unified School District, recently noted that the district's summer programs were "critical to address learning loss, provide individualized instructional support and offer unparalleled acceleration options for our students" (Harter, 2023).

It is no surprise that summer programming is an important part of districts' academic recovery efforts. Given school vacation schedules, summers offer an opportunity to provide extra support for student learning: districts can leverage existing infrastructure and resources in the summer (i.e., school buildings, buses, teachers), and families can benefit from access to childcare. The voluntary nature of summer school may also sidestep some of the political challenges associated with extending the school year (MacGillis, 2023). Pre-pandemic research has shown that well-designed summer programs increase student achievement in math (Lynch et al., 2023). The pre-pandemic impacts on reading achievement are more ambiguous: Kim and Quinn's (2013) meta-analysis of programs using research-based reading curricula found positive impacts of classroom-based interventions, but Augustine and colleagues' (2016) multisite randomized controlled trial (RCT) of district-led summer learning programs found improvements to math achievement following summer school, but no improvements to reading. Research has also shown that poor attendance can reduce program effectiveness (Augustine et al., 2016), and many summer programs with low attendance appear to have limited effects on academic success (Kim & Quinn, 2013; Lynch et al., 2023).

Despite the modest positive impacts shown in research conducted before the pandemic, there has been little research into the effects of summer learning following the pandemic. With that in mind, we assess the academic progress of students who attended summer school following the 2021–22 school year across eight districts that collectively serve about 400,000 students. Controlling for student characteristics and spring 2022 achievement, we find an average impact of summer school participation of 0.03 standard deviation (*SD*) in math.[1] We do not find statistically significant impacts in reading. To put these results in context, at

[1] As we note in the Methods section, we attempt to minimize any selection bias in our estimates by using a selection-on-observables research design. Although we use causal language through the report (e.g., "impact"), we cannot rule out the possibility that our results are biased in an unknown direction.

the end of the 2022– 23 school year, students were still -0.2 to -0.25 *SD* behind in math in upper elementary grades (Lewis & Kuhfeld, 2023). That means that if 100% of students participated in summer programming, it would have closed about 10% of the average learning loss associated with the COVID-19 pandemic in these districts. However, when we account for the fact that in this study just 13% of students enrolled in summer school in the average district, the summer school programs we studied closed only 2% to 3% of pandemic-related learning loss. In the next section, we set the stage for the study by briefly summarizing pre-pandemic research on the design of summer programs and their estimated effects on student achievement.

2. BACKGROUND

2.1 Summer Program Design and Effects

Pre-pandemic evidence on the design of summer school programs primarily comes from formative evaluations and best-practice literature (Bell & Carrillo, 2007; Boss & Railsback, 2002; McEachin et al., 2018; McLaughlin & Pitcock, 2009). This literature suggests summer programs are typically designed to be voluntary, delivered (in part or whole) by district staff at a subset of "hub" school sites, and include educational and enrichment activities. Programs vary widely, with different eligibility and invitation rules, curricula, class sizes, amounts of daily instructional time, and overall durations.

Based on a multiyear study of summer school programs in seven districts, researchers from the RAND Corporation recommended a more consistent approach to summer school, which includes both high-quality planning and curricula, hiring the district's most effective teachers, a mix of academic (3 to 4 hours, daily) and enrichment activities, and sufficient duration (5 days per week, for 5 to 6 weeks; Schwartz et al., 2018). However, implementing these recommendations can be challenging. Even with prior guidance on program design, several programs in the RAND Corporation study failed to meet the recommended criteria.

The broadest evidence about the efficacy of summer interventions comes from two meta-analyses of subject-specific programs. A meta-analysis of 41 studies of school and at-home summer reading programs in Grades K–8 showed wide variation in effects but concluded that classroom-based summer reading programs improved reading test achievement for students from low-income households by about 0.09 *SD* (Kim and Quinn, 2013).[2] A more recent meta-analysis of 37 studies of summer math programs found similarly positive effects (+0.10 *SD*) on students' standardized math achievement across income levels (Lynch et al., 2023).[3] Other evidence on the impact of voluntary, district-run summer programs comes from the RAND Corporation study noted earlier (Schwartz et al., 2018). Along with collecting programmatic details, the

[2] Fourteen (40%) of the studies employed a regression discontinuity design or leveraged an RCT to estimate program effects on reading outcomes (Kim and Quinn, 2013).
[3] Eleven (30%) of the studies employed a regression discontinuity design or leveraged an RCT to estimate program effects on math outcomes (Lynch, An, and Mancenido, 2023).

study included an RCT that examined the effects of offering two years of voluntary classroom-based summer programming for approximately 3,000 students across five districts. The results suggested the programs improved students' math achievement after the first summer (+0.08 *SD*) but had no benefits after the second summer. The programs had no impact on reading achievement in either year. However, because the districts struggled to maintain high levels of enrollment and attendance, the study's intent-to-treat analysis likely understated the benefits to students who attended the program. In particular, in the study's first year, the researchers reported that 21% of students did not show up at all to the program, 29% had low attendance (i.e., less than 80% of the time), and only 50% had high attendance (i.e., at least 80% of the time). Attendance issues increased in the second year, with less than one-third of the original sample attending more than 80% of the program.

In a related study (Augustine et al., 2016), researchers found summer school benefits were associated with students attending at least 20 days of programming. Since student enrollment and attendance were so important to the program's impact, the RAND Corporation researchers recommended that districts invest in personalized program recruitment, set firm enrollment deadlines, have clear attendance policies (and offer incentives if possible), and establish systems for monitoring enrollment and attendance (Schwartz et al., 2018).

2.2 Summer Programs and Academic Recovery From COVID-19

Systematic evidence on how summer school may have changed during the pandemic— and with what results—is scarce. As we noted earlier, survey research suggests districts expanded summer school in response to the pandemic, with most reporting new or increased programming (Diliberti & Schwartz, 2022). Other pandemic-related evidence suggests students may be seeing more academic recovery during the summer than during the school year (Lewis & Kuhfeld, 2023). In fact, during the pandemic, students appeared to suffer less from "summer slide" than they did prior to the pandemic (Lewis & Kuhfeld, 2022). For example, relative to summer 2019, summer slide in 2022 decreased on average from -0.09 *SD* to -0.07 *SD* in reading, and from -0.20 *SD* to -0.17 *SD* in math across Grades 3–8. In short, during the pandemic, districts reported expanding summer programming and students appeared to be learning more in the summer. But it is hard to know what to make of either improvement, given the lack of careful assessments of summer school as a strategy for learning recovery from COVID-19. With that in mind, we estimate impacts of summer learning programs on achievement with the goal of highlighting how these interventions fit into the broader range of strategies that schools and districts will need to meet this moment. More specifically, we examine how summer interventions have affected academic achievement in the wake of COVID-19 by using a detailed dataset from eight districts that includes details on program characteristics, student eligibility and attendance, and academic outcomes. We contextualize these findings against recent evidence on the learning loss related to COVID-19 that remains at the end of the 2022–23 school year. The following sections explains our sample, data, and approach.

3. METHOD

3.1 Data

This report draws on data from the Road to Recovery (R2R) project, an ongoing partnership between researchers and 11 school districts that aims to provide districts with timely feedback on their academic recovery interventions. Of these districts, eight provided data to participate in the summer 2022 analysis, and these districts comprise the sample for this report.[4] These eight districts collectively enroll approximately 400,000 students across seven states. As displayed in Table 1, these districts serve higher percentages of Black and Hispanic students (56%) and students eligible for free- and reduced-price lunch (55%) relative to national averages (33 and 45%, respectively).[5]

We constructed an analytic sample of students from these eight districts that met the following criteria: a student was expected to be entering Grade 1 through Grade 8 in fall 2022, was eligible for the summer program based on district-specific criteria, and had NWEA MAP Growth scores for both the spring 2022 and fall 2022 tests in reading or math.[6] We exclude students who left the district between the spring test and fall test periods. Our combined district analytic samples include 129,721 students, with numbers ranging from 1,804 to 39,248 students across districts and subjects.

Data for this study come from three primary sources: (a) interviews about program characteristics with district leaders, (b) student-level eligibility and program participation data provided by the districts, and (c) NWEA MAP Growth assessments. We describe each of these sources in more detail below. For the purposes of this study, a summer learning program was defined as a program that the district provided for formal academic support in math and/or English language arts over the summer.[7]

Interviews

We collected qualitative data on the design and implementation of summer learning programs in fall 2022 through semistructured interviews with summer programming leaders in each district. In total, we conducted nine interviews. All interviews were conducted virtually, lasted 60 minutes, and the research team followed up via email and reviewed any documentation shared by program leaders to resolve any remaining questions. The questions focused on the key design elements of each program, including student eligibility criteria, invitation processes, program duration and intensity (i.e., hours per day), daily

[4] These districts include Alexandria City Public Schools (Virginia), Dallas Independent School District (Texas), Guilford County Schools (North Carolina), Portland Public Schools (Oregon), Richardson Independent School District (Texas), Suffern Central School District (New York), Tulsa Public Schools (Oklahoma), and one district that asked to remain anonymous. For more about R2R and related research, see CALDER (n.d.).
[5] In alignment with our agreements with each of the R2R districts, we protect districts' anonymity with respect to their results by masking district names and by being purposely ambiguous about the details of specific programs. [6] An exception to this rule is one district (District 7) with optional MAP testing and very low testing rates in spring 2022. For this district, we include spring 2022 state standardized test scores as a proxy for spring 2022 MAP scores, and we define our analytic sample based on the combined availability of state standardized test scores along with winter 2022 and fall 2022 MAP scores. Our model specification for this district includes cubic polynomials of both winter 2022 MAP and state standardized test scores. Less restrictive specifications that allow for missingness in either of these variables yield consistent results.
[7] This definition excluded some summer programs offered by R2R districts from our analysis because they were focused exclusively on enrichment activities, such as art, karate, or drama. We excluded these programs from our analysis because the districts did not consider these programs to be academic recovery programs with the explicit goal of improving student achievement.

hours of instruction in each subject, delivery mode (e.g., virtual or in-person), and staffing. We also asked about whether tutoring was offered through the summer program; if yes, we asked how students were identified to receive the tutoring, when did the tutoring happen and for how long, and what did students who did not receive tutoring do during the tutoring time. Additional questions probed about the successes and challenges the district experienced in implementing the program as designed. Notes from interviews were captured in a notes template that was shared with participants during the interview and reviewed by participants following the interview to ensure accuracy.[8] The key design elements of each program are summarized in Table 2.

Student-Level Eligibility, Participation, and Dosage

Following their program interview, each district received a student-level data request tailored to their district and programs. All requests included enrollment data, attendance data, and demographic data from the 2021–22 and 2022–23 school years. Requests also included a wide variety of summer-program-specific eligibility and participation student-level variables, such as being recommended or invited to attend summer school, multi-tiered systems of support in math and English language arts, previous state test scores, summer school enrollment status, summer school site attended, and daily program attendance. We counted students as attending summer school if they attended at least one day. We also measured the dose of the program that students received as the average number of days that students attended the program, among those who attended at all; we also considered the number of hours of math or reading instruction students received each day and the percentage of total program days attended.

NWEA MAP Growth Assessments

The achievement outcome data in this study are from the NWEA MAP Growth longitudinal student achievement database. School districts use MAP Growth assessments to monitor student achievement and growth in reading and math over the course of the school year. In most districts, the tests are administered three times each year: in the fall, winter, and spring.[9] Relative to state tests administered each spring, MAP Growth assessments are particularly well-suited for the present study because they can more narrowly isolate changes in student achievement over the summer and between spring and fall assessments between two school years, rather than from spring to spring. Relative to a fixed form test, the computer-adaptive format of the MAP Growth assessment increases precision at the high and low ends of the distribution; this increased precision is particularly pertinent in the context of the pandemic, because many students are performing below grade level.

We standardize NWEA MAP Growth scores by subject and grade level using the NWEA 2020 MAP Growth norms (Thum & Kuhfeld, 2020) that are based on a nationally representative sample of students in the pre-pandemic school years (i.e., 2015–16, 2016–17, and 2017–18). We define "grade level" as students' expected rising grade in fall 2022 based on the grade they were enrolled in as of spring 2022 to have a

[8] The interview notes template is available at https://caldercenter.org/sites/default/files/District-SR-template-summer22.pdf.
[9] Schools typically administer fall tests between August and November, winter tests between December and mid- March, and spring tests between late March and June.

consistent measure of grade across students, regardless of whether or not a student repeated or skipped a grade in the following year.[10] Normalizing the scores enables us to assess students' academic performance relative to a pre-pandemic nationwide distribution of test scores.[11]

The NWEA data also include student-level demographic data on race/ethnicity and gender. School-level demographic and enrollment data linked to the NWEA dataset are from the 2020–21 Common Core of Data collected by the National Center for Education Statistics.

4. EMPIRICAL APPROACH

4.1 Identification Strategy

We use value-added models to estimate the effect of each of the eight summer programs on MAP Growth test scores, using the previous spring as the baseline and the subsequent fall score as the outcome:

$$MAP_{math,ijt} = \beta_0 + \beta_1 SS_i + \beta_2 Other_{i,math} + \beta_3 Other_{i,ela} + \tau\, \mathbf{X}_{i,t} \times Grade_{it} + \psi\ Sch\_Grade_{ij,t-1} + \varepsilon_{ijt}.$$

In the above equation, $MAP_{math,ijt}$ denotes math achievement for student *i* in school *j* in term *t* as measured by the standardized math MAP Growth score and where *t* is fall 2022. Our main treatment variable is the SS_i term, a binary indicator equal to 1 if student *i* attended at least one day of summer school during summer 2022. The estimated effect of participating in the summer school program is therefore the estimated coefficient β_1. The term $Other_i$ is an indicator variable equal to 1 if student *I* attended other academic COVID-19 recovery interventions during the summer, such as tutoring. Three of the eight districts provided tutoring sessions during the summer, mostly at summer school sites. In those districts, we considered the receipt of tutoring and summer school as an additional intervention, distinct from summer school alone or tutoring alone. The vector $X_{i'}$ includes students' prior achievement, demographics, and missing data indicators. Specifically, we include student *i*'s baseline MAP math achievement in term *t* – 1 as a cubic polynomial, where *t* – 1 is spring 2022. We additionally include prior MAP math test scores from winter 2022 and fall 2021 as well as spring 2022 MAP reading test scores. We interact all MAP scores with a categorical variable (i.e., MAP_missflag) flagging possible combinations of missingness within each subject.[12] Whenever appropriate and available, we also included any prior tests, such as state standardized tests, which were used to identify students to be prioritized for summer programs. We also include in pretreatment student covariates in $X_{i',t-1}$ that includes, when available, a student's race/ethnicity, gender, Individualized Education Program status, English language learner status, 504 plan status, and economic disadvantage status as of term *t* – 1. Finally, we include indicators for the calendar weeks during which student *i* took each

[10] For each district, we confirmed that atypical grade progression patterns were rare and that including or excluding these students in our sample did not affect our results. Of note, grade progression/retention was not contingent on summer school attendance for any district herein.
[11] $z(Y_{igst}) = (Y_{igt} - \bar{Y}_{gt}) / SD(Y_{gt})$
[12] This is commonly referred to as the missingness-indicator method. For a recent theoretical support of this method, see Zhao and Ding (2023).

math MAP Growth test in terms *t* and *t* – 1, respectively. The entire vector is interacted by grade level to account for across-grade differences in the relationship of covariates, treatment, and the outcome. $Sch\_Grade_{ij,t-1}$ represents school-by-grade fixed effects, based on the school and grade of student *i* in spring 2022. ε_{ijt} denotes idiosyncratic error. We estimate a linear model and cluster the standard error at the school-by-grade level (Abadie et al., 2022). When reading achievement is the outcome, we reverse the reading and math subscripts in equation (1).

Our primary focus is β_1, the coefficient on the summer school treatment. In order for β_1 to identify a causal relationship, we must assume that students in the treatment group (enrolled in summer school) are randomly assigned conditional on the other covariates in the model. Our selection-on-observables research design will not identify the causal effect of summer school if students were selected into summer school for reasons not observed in our data (e.g., motivation, family resources) and correlated with subsequent achievement. But it is not possible to a priori sign the direction of any bias. For example, it is possible that students who attended summer school regularly had higher motivation or family resources (e.g., an adult available to get them to school) compared to students who did not attend regularly or at all, positively biasing treatment effects. On the other hand, families may have opted out of summer school if they needed longer childcare throughout the day or signed their students up for programs that not only were longer but also had an educational benefit, negatively biasing the estimates. That said, there are two reasons to believe the results capture the true effect of summer school. First, researchers have used value-added approaches in program evaluation to isolate the causal effect of teachers and schools (Abdulkadiroglu et al., 2011; Bacher-Hicks et al., 2019; Chetty et al., 2014; Deming, 2014). Second, as we describe below, given the length of the summer school programs in our study, the positive effects in math we observe are not far off from what would have been expected based on pre-pandemic experimental estimates of the impact of summer school attendance.

We adapt the above general specification based on each district's available data and summer school program design. For districts that used data and/or created decision rules that combined multiple data sources to target or prioritize students for participation (e.g., scores below a certain threshold on MAP assessments, state tests, and/or other academic assessments), we control for a binary measure of prioritization in addition to any other indicators or test scores (in cubic form) related to prioritization. As with the MAP scores described above, we impute missing values of these additional achievement measures and interact them with imputation flags and expected rising grade level.

In addition to estimating the effect of summer school participation, we estimate the effect of an hour of math (or reading) instruction during summer school on fall 2022 MAP scores. For these specifications, we substitute binary measures of participation in summer school and/or other academic COVID-19 recovery interventions taking place during the summer with continuous measures of the number of hours of instruction received in these programs. In most cases, these measures are based on daily attendance data received from the district and information from district interviews on the intended hours of subject-level instruction per day.

Last, we synthesize our estimates of summer program effects from the eight districts by conducting a meta-analysis. For each subject, we estimate meta-analytic estimates with a random effect model with restricted maximum likelihood (REML) estimation (DerSimonian & Laird 1986; Hedges, 1983; Raudenbush, 2009).

4.2 Contextualizing Our Estimates

We use the estimated effects from the pre-pandemic summer school impacts literature to help contextualize our findings. Specifically, we compare our estimates to those we would expect to see, given the observed dosage (in days) of summer school and benchmark effect sizes found in the existing literature on the academic impacts of summer school. For math, we use meta-analysis estimate of 0.101 *SD* (see Lynch et al., 2023), based on summer programs averaging 5.2 weeks (i.e., 26 days) in length, with an average of 2.1 hours of math instruction. For reading, we use a meta-analysis estimate of 0.09 *SD* for classroom-based summer interventions (see Kim & Quinn, 2013), and we assume that the program length and daily dosage of reading instruction were similar to those of the math benchmark estimate. Assuming a linear relationship between the amount of time spent in instruction during summer school and the total average effect of the program, we back out an hourly expected effect of summer school instruction (e.g., 0.101 *SD*/(26*2.1) *SD* in math), which we then multiply by the average dosage of instruction received in each subject to arrive at the "expected effect" of summer school participation in each district.

We similarly calculate an "expected effect" for interventions that combined summer learning programs with tutoring. For these cases, which occur in three of our districts, we also use findings from a meta-analysis of studies on high-quality, high-dosage tutoring programs (Nickow et al., 2020) that estimate gains of 0.38 *SD* (0.35 *SD*) in math (reading) achievement assuming three tutoring sessions per week. Over a 36-week school year, that equals 108 hours of tutoring. To arrive at the total expected effect for this combined intervention, we add the expected effect from the average dosage of summer school instruction with the expected effect from the average dosage of additional tutoring instruction.

5. RESULTS

5.1 Participation and Dosage

The percentage of students who attended at least one day of summer programming for each of the eight R2R districts in the sample is seen in Table 3, Column 4.[13] In the districts where we could observe attendance (and not simply summer school sign-ups), the participation rates varied substantially across the districts, ranging from 4.8% to 22.6%. The one exception was District 3, in which 46.2% of students enrolled but

[13] It is important to note that students in some grade levels were omitted from the analytic sample due to sparse MAP score availability (which varied by district), even though they were offered summer programming. If we include these grades, participation rates in some districts were slightly higher than described in this study. For example, while the participation rate for District 8 was 4.8% for rising students in Grade 6 to Grade 8, the rate across all rising students in first through ninth grade (as the program was designed), the participation rate was slightly higher.

an unknown percentage of students attended. The average participation rate across the seven districts is approximately 13.1%, similar to the proportion of households with children under age 18 that reported children's attendance in summer learning programs in summer 2022 (approximately 10%) in the nationally representative 2022 Household Pulse Survey data.[14] Across the R2R districts, average participation rates were higher among students whose schools were host ("hub") sites for the summer program (17.3%) relative to students whose schools were not summer sites (11.9%).

Most districts targeted recruitment for their summer programs toward students they considered to be most in need of additional support, but also allowed (and often encouraged) nontargeted students to attend if they wanted. More specifically, six of the eight R2R districts (Districts 2, 3, 4, 6, 7, and 8) targeted recruitment efforts for their summer programs at subsets of struggling students using criteria that ranged from failing to meet threshold scores on state standardized tests to the discretion of classroom teachers. The resulting groups of prioritized students comprised 28% to 55% of the districts' populations, and participation rates were higher among these students than nontargeted students across districts that used targeting (see Appendix Table A1). In District 7, all the students who participated were students who were prioritized for enrollment. The higher proportions of students attending summer programs in districts with targeted r ecruitment aligns with additional evidence in Appendix Table A1 that targeted students who participated in summer programs at these districts tended to score substantially lower on the MAP Growth tests, by about 17 Rasch UnIT (RIT) points in math and 11 RIT points in reading, prior to the summer. Participation rates in these districts were also higher for students in disadvantaged subgroups, including students receiving special education services, students who are English language learners, students who are economically disadvantaged, and students who are Black or Hispanic. At the two districts (Districts 1 and 5) that did not use targeting and encouraged all students to attend, participation rates varied less by spring MAP scores (participants scored 3 RIT points lower in math and 5 RIT points lower in reading) and across student subgroups. (See Appendix Table A1 for a breakdown of participation rates by subgroups.)

Notably, District 1 had the highest rate of participation within our sample. A distinct design feature that likely drove this participation rate was District 1's extended summer program hours (see Table 2), which operated from 8 a.m. to 5:30 p.m. for elementary school grades and from 9 a.m. to 4 p.m. for secondary school grades, thus providing child care for working parents. Moreover, the district leaders expressed that framing their summer programs as "summer camp"—an exciting learning and enrichment program during the summer—rather than as "summer school"—may have helped recruited students to opt into the program.

As we show in Table 2, the intensity of summer programs varied across districts: summer programs offered between 15 and 20 days of programming, with anywhere between 45 minutes and 2 hours of daily

---

[14] We use data from the Week 49 Household Pulse Survey Public Use File release (https://www.census.gov/programs-surveys/household-pulse-survey/datasets.html) and limit the sample to respondents with children under age 18 enrolled in public schools.
pected effect from research. Two other districts (3 and 5) had positive though not statistically significant estimates, though sample sizes in those districts were also notably small.

academic instructional time per subject (i.e., math and reading). Six districts convened four days per week, and the other two districts offered five days per week of programming. Total hours of academic instruction ranged from 23 to 67 hours across programs. Overall, these planned dosages are notably smaller than the recommended minimum of 75 hours of academic instruction per program (see Schwartz et al., 2018), and are smaller than those of summer programs evaluated in the previously mentioned studies and that have been evaluated and documented in the literature. For instance, five district-led summer programs evaluated by the National Summer Learning Project (Augustine et al., 2016; McCombs et al., 2014) offered between 23 and 29 days of instruction. Similarly, the average length of summer programs reviewed by Lynch and colleagues (2023) was approximately 5.2 weeks (26 days, assuming five days per week of programming).

Of note, three districts (Districts 2, 4, and 6) offered tutoring to a subset of students during their summer programs. District 6 designed their program with the goal of delivering two to three 30-minute tutoring sessions per week during the time that students would otherwise be participating in enrichment activities, effectively providing three to seven supplemental hours of instructional time beyond what was provided through the standard program. However, students in Districts 2 and 4 were generally pulled out of their academic summer program classes to receive tutoring, such that the tutoring hours effectively substituted for other academic time and did not increase the overall dosage of instruction students received through the program. Attendance plays a key role in the ability for summer programming to affect student outcomes, and we see evidence of significant heterogeneity across districts in student participation rates. Table 3 (Column % Total Days Attended) shows that the proportion of days attended relative to the number of intended days varied from 58% to 80%, with the average of 68.5%. Across the districts, this translates to students attending between 9.9 and 13.6 days of summer school, and— accounting for attendance rates—receiving 60 to 120 minutes per day of instruction per subject and approximately 14 to 27 total hours of summer instructional time per subject.[15]

5.2 Impact Estimates

Academic effects of summer school

Figures 1 and 2 show the estimated effects of attending at least one day of summer school on fall MAP Growth 2022 math and reading scores, respectively, across all the grade levels examined, as well as the expected effect based on the average dosage received in each district. We focus on the effect of attending at least one day of summer school because it is difficult to account for selection biases that lead to differential uptake of treatment in observational data. At the top of both figures, the "overall" estimate reflects results from a cross-district meta-analysis using a restricted maximum likelihood (REML) model. While the analytic samples and model specifications vary somewhat across districts because of summer program design and data availability, data limitations in two districts are worth noting. In District 3, we have data only on whether a student was enrolled in the summer program, not whether they attended, and as such the point

[15] Proportion of scheduled days students attended in District 4 was low at 58%. The summer program was separated into two blocks, with a one-week break in-between, which likely affected the number of days students attended.

estimate shown is of the estimated treatment effect of simply being enrolled.[16] In District 8, representative MAP Growth scores are available only in middle school grades, limiting our analysis to these grades despite that elementary students participated in the summer program as well. Given these distinctions, we also show an overall estimated effect that excludes Districts 3 and 8.

While the estimates across districts vary, outside of Districts 7 and 8, all are positive for math test scores and quite close to the aggregated (across districts) point estimates of roughly 0.03. This aggregated figure (0.032 when Districts 3 and 8 are excluded) is also nearly equivalent to the expected effect from research (0.036 *SD*).[17]

Figure 3, which shows the estimated effect of each hour of summer school instruction in math, offers similarly positive results. The overall estimate excluding Districts 3 and 8 is 0.0018 *SD*, equivalent to the expected hourly effect (0.0018 *SD*). Five of the eight districts (1, 2, 4, 5, and 6) have statistically significant positive hourly effects, ranging from 0.0014 *SD* to 0.0040 *SD*.

Our findings are less promising when it comes to effects on reading scores. As shown in Figure 2, the effect of attending at least one day of summer school was positive and statistically significant only in District 6, where the estimated effect (0.029 *SD*) was roughly equivalent to the expected effect. In the remaining districts, the estimated effects are statistically indistinguishable from zero. Similarly, estimates of hourly effects of summer school on reading, shown in Figure 4, are statistically indistinguishable from zero in all districts except for District 6 (0.0013 *SD*, where the estimate was indistinguishable from the expected hourly effect of 0.0017 *SD*). While overall these results are smaller in magnitude than what we would expect,[18] they are consistent with the findings from a multisite RCT of district-led summer learning programs conducted by RAND Corporation (Augustine et al., 2016), which found improvements to math achievement following summer school but not to reading.

Subgroup analyses

In addition to estimating the effects of summer programming across all grade levels included in our analysis, we examine effects by elementary and middle school grade ranges. The results of these grade range analyses are shown in Tables 4 and 5 for math and reading, respectively. When broken out by grade range, we find that younger students—those rising in Grade 5 and below—drive the positive gains in math. In nearly every district, the magnitude of the estimated effect is considerably greater among elementary grades than middle school grades, with the exception of District 5, where the estimated effects are equivalent. By and large, the program length and attendance rates among participants were similar across grade ranges within each district, suggesting that the differential effectiveness was not a result of different amounts of dosage.

[16] District 3's estimate would therefore be negatively biased if interpreted as the impact of attending at least one day of summer school.
[17] Four of the eight districts (1, 2, 4, and 6) had statistically significant estimated effects, all of which were close to or exceeded the expected effect from research. Two other districts (3 and 5) had positive though not statistically significant estimates, though sample sizes in those districts were also notably small.
[18] Based on Kim and Quinn (2013), whose meta-analysis focused on programs using research-based reading curriculums.

We also estimated the effects of summer school by different student subgroups, including by race and ethnicity, free-or-reduced-price lunch status, special education status, English learner status, prior performance levels, and by eligibility for targeting for enrollment in summer school (See Appendix Tables A2-A8).[19] Across all districts, we found little evidence for heterogeneity of impact across these subgroups, though, in many cases, small sample sizes limited our ability to detect differences across groups. In District 6, one of our larger districts, we detected strong positive gains in math overall, and we found significant positive gains particularly among Black students, who comprised approximately 53% of summer school participants, relative to 41% of the overall student population.

Academic effects of summer school with tutoring

Figure 5 shows the estimated effects in math and reading of attending any amount of summer school that replaced or supplemented classroom instruction with tutoring. Our estimated impacts of these programs are similar to our estimated impacts for receiving the standard summer learning program in each of the three districts that offered summer school with tutoring. Among these districts, only Districts 2 and 4 (where tutoring time typically replaced classroom instruction) have positive and statistically significant effects of the combined program on math (0.038 *SD* [District 2] and 0.047 *SD* [District 4]), though the (nonsignificant) point estimate for District 6 (where tutoring time supplemented classroom instruction) is of similar magnitude (0.044 *SD*). It is worth noting that fewer students participated in combined summer school with tutoring than in summer school alone. The total number of students taking part in this intervention for math was 2,477 in District 2; 826 in District 4; and 790 in District 6 (in reading, 1,686, 1,096, and 425, respectively).[20] Our statistical power for this analysis was therefore limited, particularly in District 6, where only 2.4% of the student population in the relevant grade levels participated, suggesting that the nonsignificance of its estimated effect is likely driven in part by sample size. In reading, on the other hand, none of the estimated effects are statistically distinguishable from zero, and the magnitudes are considerably smaller (or negative) relative to math. The effect of the standard summer learning program is statistically equivalent to summer school with the addition of tutoring.

6. DISCUSSION AND CONCLUSIONS

Across the eight districts in the study, we found that the structure (e.g., size relative to district enrollment, number of days, hours per day) of summer programs and take-up rates varied substantially. Students who attended summer programs tended to make small but significant improvements in math achievement, but no significant gains on reading tests relative to similar peers who did not attend. Our results are important not only for adding additional evidence to the use of summer programs to improve student learning but also whether these programs can make noticeable headway into the learning losses caused by the COVID-19 pandemic.

We add to the existing evidence on summer learning programs impacts on reading and math achievement.

[19] We do not present results by subgroup for District 5 due to sample size.
[20] These participant numbers are among students in the analytical sample of each district, meaning that they are in the relevant grade levels and that they have valid spring and fall 2022 MAP scores.

For reading, we estimate null effects of attending a program in all but one district and can rule out even small positive effects of at least 0.03 *SD* for most programs. The general lack of effects on reading aligns with the null reading effects estimated in the RCT study by Schwartz et al. (2018). The results are disappointing when compared to the average reading achievement effects of classroom-based summer reading interventions estimated in the meta-analysis of Kim and Quinn (2013). We speculate that one reason it may be more difficult to achieve reading gains than math gains for summer school participants relative to nonparticipants is because nonparticipants may also practice reading over the summer but may be less likely to practice math. This explanation would align with evidence that shows larger effects of school inputs on math achievement than reading (e.g., Burgess et al., 2023; Riehl & Welch, 2022). To inform the interpretation of the impact of summer programs on reading achievement, future research should examine the summer experiences of students who were eligible or invited to attend but did not participate.

Our analysis of math gains, by contrast, found participants' spring to fall test score gains per hour of academic programming were positive and remarkably similar across districts. They were also similar to what would be expected based on the pre-pandemic literature at six of the eight R2R districts (Lynch et al., 2023); perhaps surprisingly, the estimated math gains were prevalent even in districts that served larger proportions of their students, up to 23% of their students in rising Grade 3 to Grade 8. These broad-based findings suggest that adding additional days of programming (with the same or more instructional time) could yield additional math gains for students. The consistency in math gains across different student subgroups indicates that, when summer program enrollment is limited, increasing the targeted recruitment and attendance of struggling students may be an effective strategy for boosting achievement among students with the greatest academic needs. Indeed, we find that low-performing and disadvantaged students comprised higher proportions of summer programs that targeted struggling students than programs that encouraged all students to participate. That said, districts that targeted recruitment still were far from reaching all of the prioritized students: on average, only 25% of targeted students participated. As the American Rescue Plan Elementary and Secondary School Emergency Relief funding lapses and districts' resources for summer programming likely decline, districts should consider replacing open enrollment policies with more targeted efforts at recruiting lower achieving students. Assigning schools that serve higher proportions of these students to be hub sites for the summer could be one such effective strategy, as participation rates were consistently higher among students whose school sites hosted summer programming than those whose did not.

To put our overall findings in context, we consider the magnitude of the summer school math gains (but not reading because we did not find significant effects) against the scale of the learning loss related to COVID-19. On the National Assessment of Educational Progress, student achievement declined from 2019 to 2022 by 0.16 *SD* and 0.20 *SD* in Grades 4 and 8 math, respectively, with greater declines in high-poverty districts; districts with high percentages of students who are Black, Hispanic, and Native American; and districts that spent more time in remote or hybrid instruction during the 2020–21 school year (Fahle et al., 2023). More recent evidence—based on MAP assessments indicates recovery during both the 2021–22 and 2022–23 school years—was minimal, with students in Grades 3–8 still lagging their pre-pandemic peers by 0.16 *SD* to 0.27 *SD*, respectively by year, only slightly improved (and worse in two grades) from 0.22 *SD* to 0.26 *SD* losses, respectively by year, estimated from fall 2021 (Goldhaber, Kane, McEachin, & Morton, 2022;

Goldhaber, Kane, McEachin, Morton, Patterson, et al., 2022; Lewis & Kuhfeld, 2023). Consistent with our findings that summer programs were associated with math gains, the national research shows summer slide in math during 2022 was smaller than typical, accounting for much of the limited academic recovery seen from fall 2021 to fall 2022 (Lewis & Kufeld, 2022; Quinn & Polikoff, 2017).

But, a simple back-of-the-envelope calculation suggests that the summer interventions we investigated resulted in very small math gains relative to the scale of districts' remaining academic recovery in math. Specifically, multiplying the weighted average effect of summer programs on student achievement, 0.03 *SD*, by the average share of students attending them, about 13%, yields an impact of 0.0039 *SD* on district-level achievement. This is roughly 2% to 3% of the overall magnitude of pandemic-related learning loss in math. Even if districts facing COVID-19–related learning losses of 0.2 *SD* to 0.3 *SD* delivered best-case scenarios in summer school programs (i.e., longer than five weeks, with more than two hours of daily math instruction; or metrics that boost math gains by 0.10 *SD*), they would need to send every student to summer school for two to three years in a row to get back to pre-pandemic math test levels (Lynch et al., 2023). The limited participation in and duration of the summer learning programs documented herein suggests that summer school, especially if expanded, can be a valuable part of a district's academic recovery strategy for math, but it must be only one of many other repeated interventions and supports if we are to make substantial progress toward recovery for all students.

Although the math gains show summer programming can be an effective district-led strategy for boosting math achievement, one or more years of programming at this scale will not be near enough for most districts to reach academic recovery from COVID-19. Alarmingly, most parents and families are unaware of how far their children have fallen behind and report they are not concerned about learning loss (Polikoff & Houston, 2022). Providing parents with accurate and accessible information about their students' academic progress needs to be a top priority for school districts and states. In the meantime, school districts and states need to continue to drive the expansion of interventions that supplement instructional time, including summer school but also high-dosage tutoring, double-dose math courses, extended school days or years, and evidence-based retention programs. As ESSER winds down and efforts to expand learning time meet resistance, district recovery efforts will also need ongoing resources and political cover. Our findings here and elsewhere (Carbonari et al., 2022) underscore the need, up and down the system, for a continued commitment to delivering recovery interventions at the scale and intensity needed to address the pandemic's academic impact. Failing to do so will have dire consequences for many students and society.

REFERENCES

Abadie, A., Athey, S, Imbens, G. W., Wooldridge, J. M. (2022). When should you adjust standard errors for clustering? *Quarterly Journal of Economics*, 138(1), 1–35. https://doi.org/10.1093/qje/qjac038

Abdulkadiroglu, A., Angrist, J. D., Dynarski, S., Kane, T., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics*, 126, 699–748.

Augustine, C. H., McCombs, J. S., Pane, J. F., Schwartz, H. L., Schweig, J., McEachin, A., & Siler-Evans, K. (2016). *Learning from summer: Effects of voluntary summer learning programs on low-income urban youth*. RAND Corporation.

Bacher-Hicks, A., Chin, M., Kane, T. J., & Staiger, D. O. (2019). An experimental evaluation of three teacher quality measures: Value-added, classroom observations, and student surveys, *Economics of Education Review,* 73, 101919.

Bell, S. R., & Carrillo, N. (2007) Characteristics of effective summer learning programs in practice. *New Directions for Youth Development* no. 114, 45–63.

Boss, S., & Railsback, J. (2002). *Summer school programs: A look at the research, implications for practice, and program sampler.* Northwest Regional Educational Laboratory.

Burgess, S., Rawal, S., & Taylor, E. S. (2023). Teachers' use of class time and student achievement. *Economics of Education Review,* 94.

CALDER (n.d.) T*he Road to COVID Recovery: Actionable Research on District Strategies for Student Advancement.* https://caldercenter.org/covid-recovery

Carbonari, M. V., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Hashim, A., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2022). *The Challenges of Implementing Academic COVID Recovery Interventions: Evidence from the Road to Recovery Project*. CALDER Working Paper No. 275-1222.

Chetty, Raj, Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104 (9), 2593–2632.

Deming, D. J. (2014, January). *Using school choice lotteries to test measures of school effectiveness* (NBER Working Paper No. 19803). The National Bureau of Economic Research.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials,* 7(3), 177–188.

https://doi.org/10.1016/0197-2456(86)90046-2

Diliberti, M. K., & Schwartz, H. L. (2022). *Districts continue to struggle with staffing, political polarization, and unfinished instruction: Selected findings from the fifth American School District Panel Survey*. RAND Corporation. https://doi.org/10.7249/RRA956-13

DiMarco, B., & Jordan, P. W. (2022a). School districts' post-Covid strategies for summer learning. *FutureEd Explainer,* May 2. https://www.future-ed.org/school-districts-post- covid-summer-learning-strategies/

DiMarco, B., & Jordan, P. W. (2022b). How local educators plan to spend billions in federal COVID aid. *FutureEd Explainer*, June 7. https://www.future-ed.org/local-covid-relief- spending/

Fahle, E. M., Kane, T. J., Patterson, T., Reardon, S. F., Staiger, D. O., & Stuart, E. A. (2023). *School district and community factors associated with learning loss during the COVID- 19 pandemic*. Center for Education Policy Research, Harvard University. https://cepr.harvard.edu/sites/hwpi.harvard.edu/files/cepr/files/explaining_covid_losses_5 .23.pdf

Goldhaber, D., Kane, T. J., McEachin, A., & Morton, E. (2022). *A comprehensive picture of achievement across the COVID-19 pandemic years: Examining variation in test levels and growth across districts, schools, grades, and students.* CALDER Working Paper No. 266-0522.

Goldhaber, D., Kane, T. K., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2022). *The consequences of remote and hybrid instruction during the pandemic.* CALDER Working Paper No. 267-0522.

Harter, C. (2023). This summer LAUSD tackles learning loss in a fun-filled environment. *Los Angeles Daily News,* June 26. https://www.dailynews.com/2023/06/26/this-summer- lausd-tackles-learning-loss-in-a-fun-filled-environment/

Hedges, L. V. (1983). Combining independent estimators in research synthesis. *British Journal of Mathematical and Statistical Psychology*, 36, 123–131. https://doi.org/10.1111/j.2044- 8317.1983.tb00768.x

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from Kindergarten to Grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, 83(3), 386–431.

Lewis, K., & Kufeld, M. (2022). *Progress toward pandemic recovery: Continued signs of rebounding achievement at the start of the 2022–23 school year.* [Research brief.] NWEA.

Lewis, K., & Kuhfeld, M. (2023). Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery. NWEA.

Lewis, K., Kuhfeld, M., Ruzek, E., & McEachin, A. (2021). *Learning during COVID-19: Reading and math achievement in the 2020–21 school year.* [Research brief.] NWEA.

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

MacGillis, A. (2023). What can we do about pandemic-related learning loss? *The New Yorker*, June 19. https://www.newyorker.com/magazine/2023/06/26/what-can-we-do-about- pandemic-related-learning-loss

McCombs, J. S., Pane, J. F., Augustine, C. H., Schwartz, H. L., Martorell, P., & Zakaras, L. (2014). *Ready for fall? Near-term effects of voluntary summer learning programs on low- income students' learning opportunities and outcomes.* RAND Corporation. https://doi.org/10.7249/RR815

McEachin, A., Augustine, C. H., & McCombs, J. (2018). Effective summer programming: What educators and policymakers should know. *American Educator,* 14(1), 10–11.

McLaughlin, B., & Pitcock, S. (2009). *Building quality in summer learning programs: Approaches and recommendations.* Wallace Foundation.

Nickow, A., Oreopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on preK-12 learning: A systematic review and meta-analysis of the experimental evidence* (Working Paper No. 27476). National Bureau of Economic Research.

Polikoff, M., & Houston, D. (2022, September). Experts Say Kids Are Far Behind After COVID; Parents Shrug. Why the Disconnect? *The74m.* https://www.the74million.org/article/experts-say-kids-are-far-behind-after-covidparents- shrug-why-the-disconnect/

Quinn, D. M., & Polikoff, M. (2017). Summer learning loss: What is it, and what can we do about it? Brookings. https://www.brookings.edu/articles/summer-learning-loss-what-is-it- and-what-can-we-do-about-it/RAND Corporation. (2016).

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta- Analysis* (pp. 295–315). Russell Sage Foundation.

Riehl, E., & Welch, M. (2022). Accountability, test prep incentives, and the design of math and English exams. *Journal of Policy Analysis and Management* 42(1), 60–96.

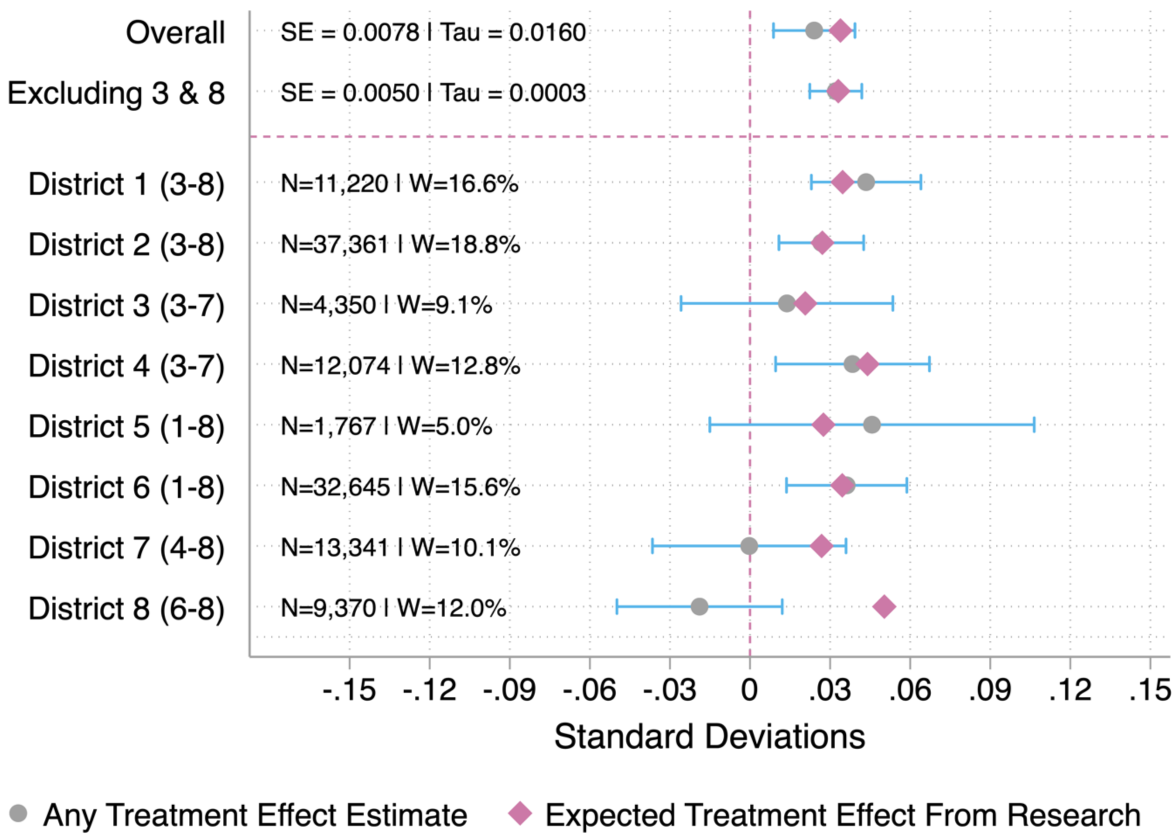
Schwartz, J. L., McCombs., J. S., Augustine, C. H., & Leschitz, J. T. (2018). *Getting to work on summer learning: Recommended practices for success*, 2nd ed. RAND Corporation.

Thum, Y.M. & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth achievement status and growth norms for students and schools.* NWEA.

Zhao, A., & Ding, P. (2022). To adjust or not to adjust? Estimating the average treatment effect in randomized experiments with missing covariates. *Journal of the American Statistical Association*, 1–11.ps://teach.map-nwea.org/impl/NormsTables.pdf.

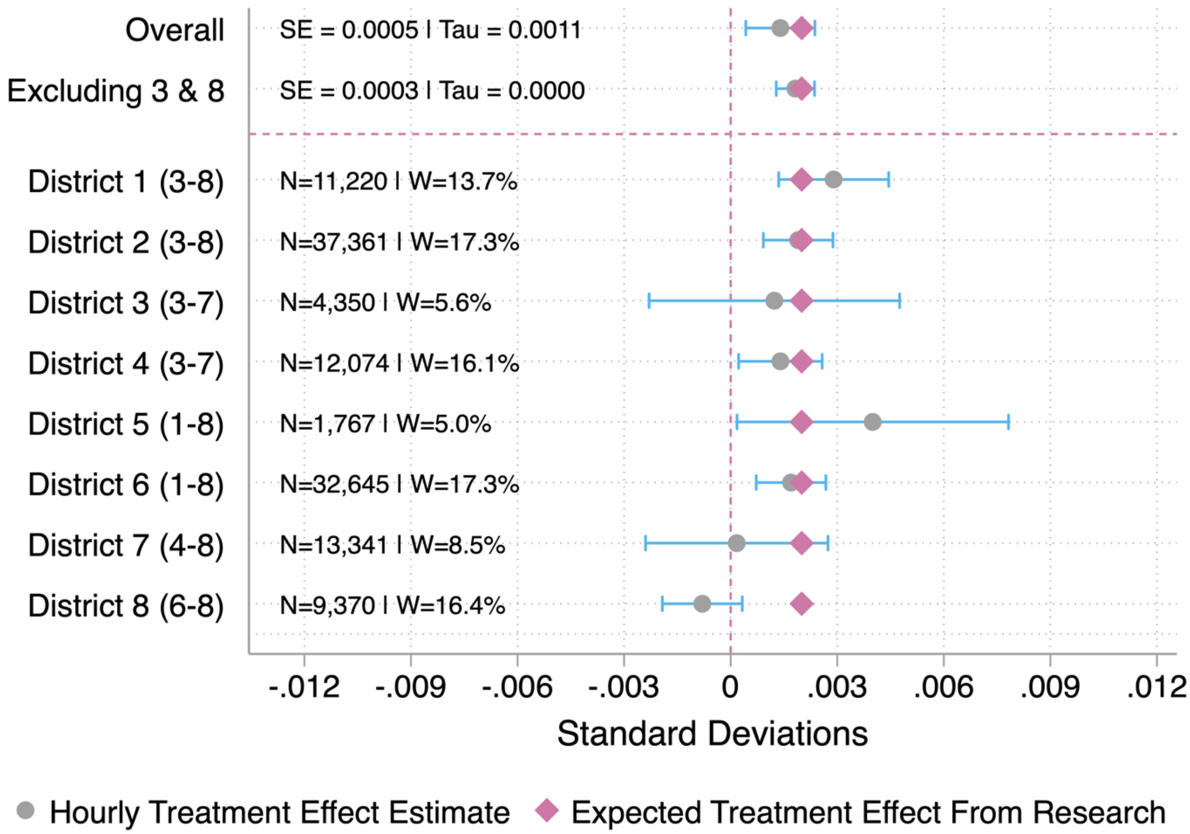
FIGURES AND TABLES

Figure 1. Summer Programs: Effect of Treatment on Math Outcomes



Notes:
1. Expected effect from research is based on a 0.101 SD increase in math achievement over 26 days of summer school with 2 hours of academic instruction per day.
2. N indicates the number of students in the analysis. W indicates the weight given to the analysis.
3. Overall estimates are constructed using a random effects model with a REML estimation. SE indicates the standard error which comes from each estimate's weighted errors, while Tau indicates the error arising from inter-estimate heterogeneity.
4. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
5. For District 8, the analysis sample only consisted of middle school grades with low participation rates, which tended to have lower effects overall. See Grades 6-8 subgroup estimates.

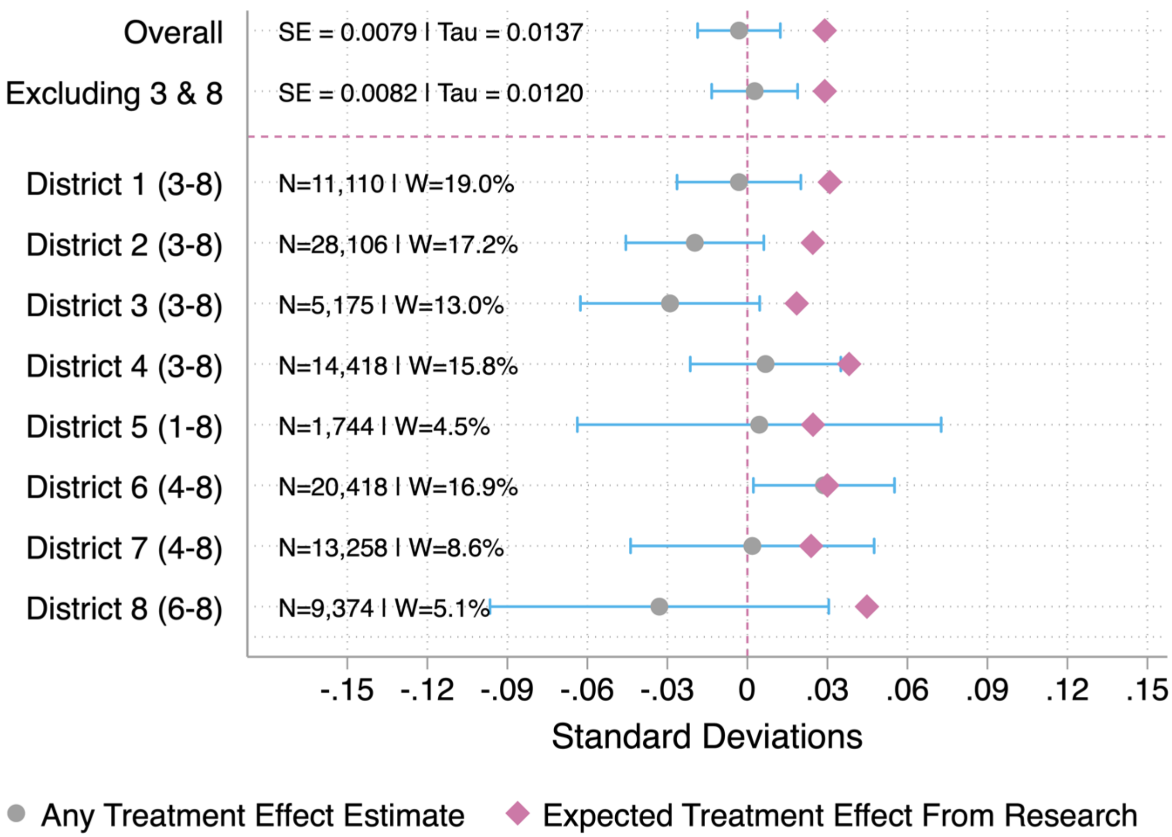
Figure 2. Summer Programs: Hourly Effect of Treatment on Math Outcomes



Notes

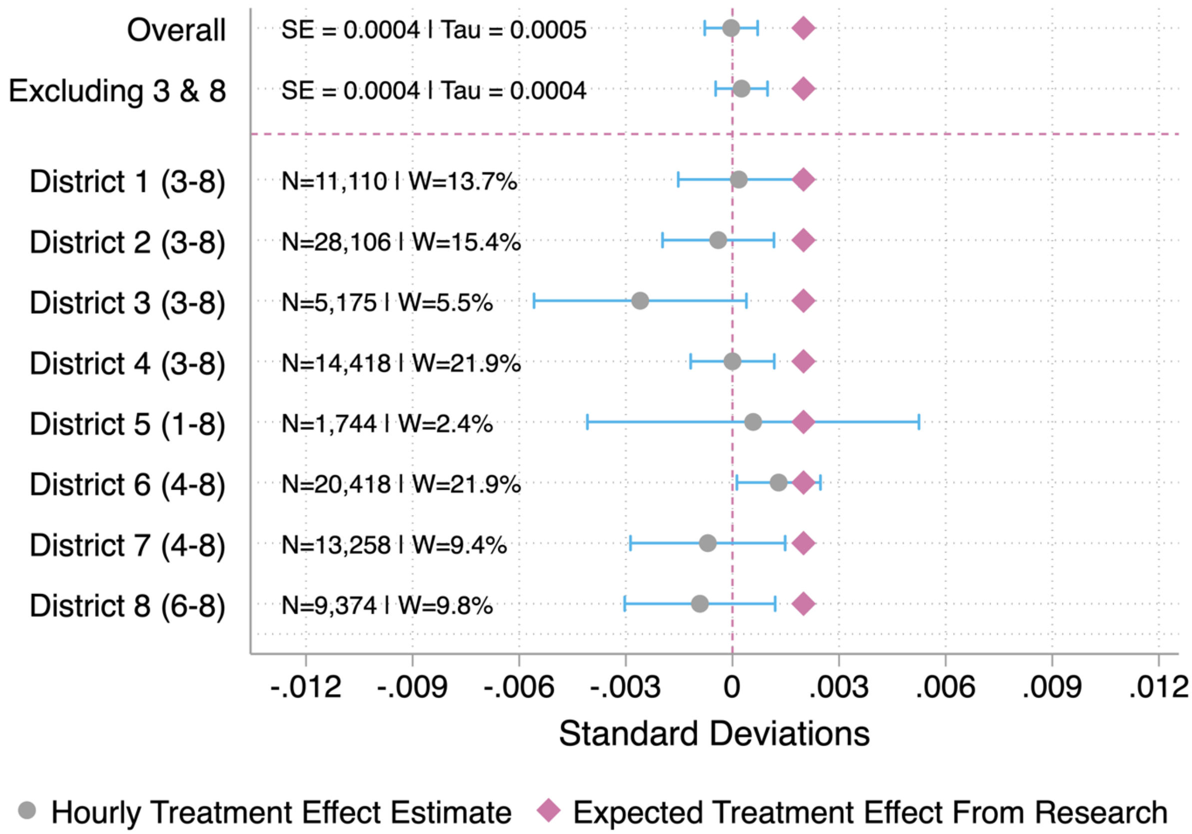
1. Expected effect from research is based on a 0.101 SD increase in math achievement over 26 days of summer school with 2 hours of academic instruction per day.
2. N indicates the number of students in the analysis. W indicates the weight given to the analysis.
3. Overall estimates are constructed using a random effects model with a REML estimation. SE indicates the standard error which comes from each estimate's weighted errors, while Tau indicates the error arising from inter-estimate heterogeneity.
4. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
5. For District 8, the analysis sample only consisted of middle school grades with low participation rates, which tended to have lower effects overall. See Grades 6-8 subgroup estimates.

Figure 3. Summer Programs: Effect of Treatment on Reading Outcomes



Notes:
1. Expected effect from research is based on a 0.09 SD increase in reading achievement over 26 days of summer school with 2 hours of academic instruction per day.
2. N indicates the number of students in the analysis. W indicates the weight given to the analysis.
3. Overall estimates are constructed using a random effects model with a REML estimation. SE indicates the standard error which comes from each estimate's weighted errors, while Tau indicates the error arising from inter-estimate heterogeneity.
4. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
5. For District 8, the analysis sample only consisted of middle school grades with low participation rates, which tended to have lower effects overall. See Grades 6-8 subgroup estimates.

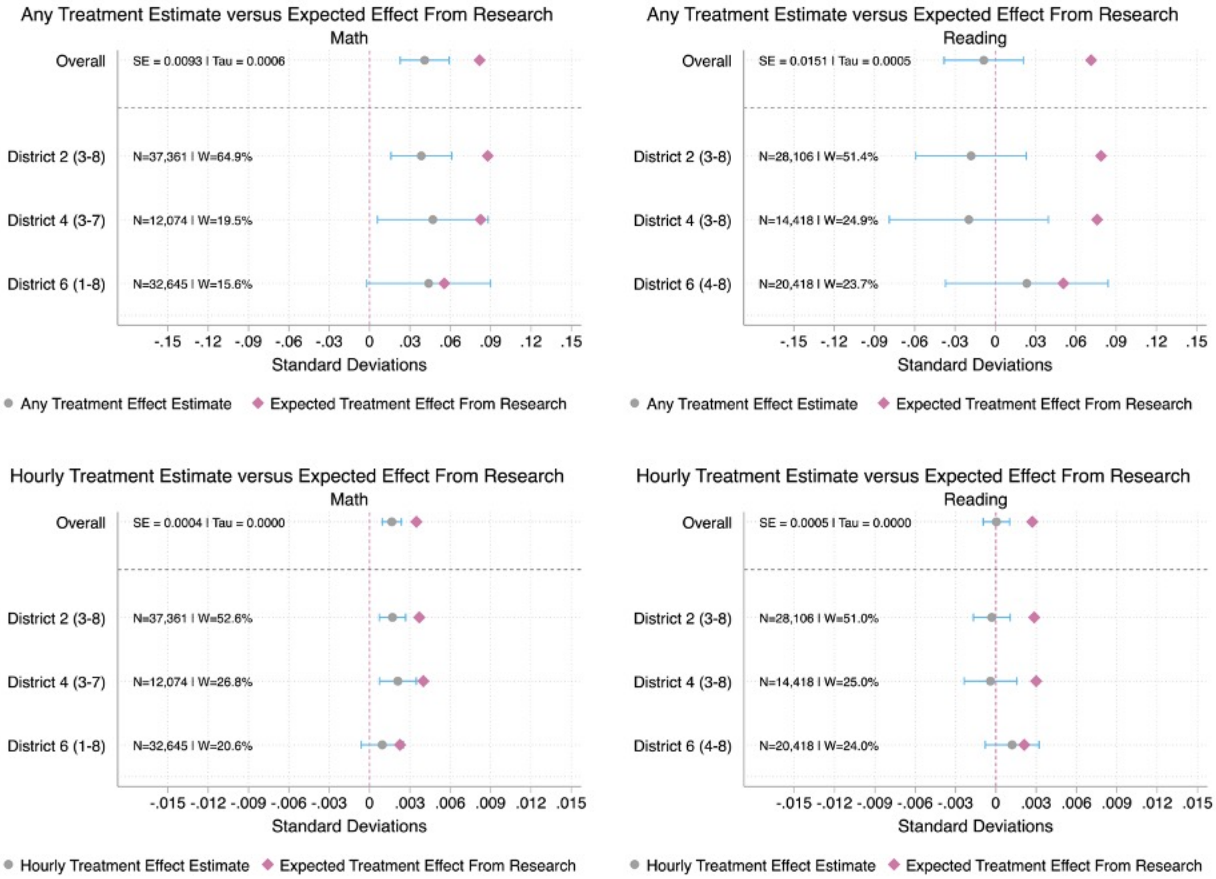
Figure 4. Summer Programs: Hourly Effect of Treatment on Reading Outcomes



Notes:

1. Expected effect from research is based on a 0.09 SD increase in reading achievement over 26 days of summer school with 2 hours of academic instruction per day.
2. N indicates the number of students in the analysis. W indicates the weight given to the analysis.
3. Overall estimates are constructed using a random effects model with a REML estimation. SE indicates the standard error which comes from each estimate's weighted errors, while Tau indicates the error arising from inter-estimate heterogeneity.
4. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
5. For District 8, the analysis sample only consisted of middle school grades with low participation rates, which tended to have lower effects overall. See Grades 6-8 subgroup estimates.

Figure 5. Effects of Summer Programs and Tutoring



Notes:

1. For math, expected effect is based on a 0.101 *SD* increase in math achievement over 26 days of summer school with 2 hours of academic instruction per day. For reading, expected effect is based on a 0.09 *SD* increase in reading achievement over 26 days of summer school with 2 hours of academic instruction per day.
2. *N* indicates the number of students in the analysis. *W* indicates the weight given to the analysis.
3. Overall estimates are constructed using a random effects model with a restricted maximum likelihood (REML) estimation. SE indicates the standard error that comes from each estimate's weighted errors, while Tau indicates the error arising from inter-estimate heterogeneity.

Table 1. Sample Demographics

	R2R Districts	U.S. Public Schools
Average district enrollment	50084	2674
Average school enrollment	678	497
FRPL-eligible (%)	55.2%	45.4%
Race (%)		
Asian	4.7%	2.9%
Hispanic	32.3%	20.1%
Black	23.2%	12.7%
White	33.5%	59.4%
School locale (%)		
City	88.47%	18.5%
Suburb	29.26%	22.7%
Town	0%	14.2%
Rural	8.59%	42.1%

Note: FRPL=free or reduced priced lunch. Data are from the Common Core of Data (CCD) collected by the National Center for Education Statistics during the 2020-21 school year.

Table 2. Designs of Summer Programs

	District 1	District 2	District 3	District 4	District 5	District 6	District 7	District 8
Subject	Math and ELA	Math and ELA	Math and ELA	Math and ELA	Math and ELA	Math and ELA	Math and ELA	Math and ELA
Grade Level	Rising K-12	Rising PK-8	Rising K-8	Rising PK-8	Rising 1-8	Rising K-12	Rising 1-8	Rising 1-9
Eligible Schools	All schools	Reg. calendar schools	All schools	All schools	All schools	All schools	All schools	All schools
Participation: Opt-in or By Invitation?	Opt-in	Opt-in and by invitation	Opt-in and by invitation	Opt-in and by invitation	Opt-in	By invitation	Opt-in and by invitation	Opt-in and by invitation
Invitation: To Whom?	None	Students scoring below grade-by-subject threshold on state tests	Students with academic, SEL, or other needs, based on own prioritization matrix	Students scoring below grade-by-subject threshold on state or MAP tests	None	Low-scoring students	Low-performing and historically underserved students	Students identified as academically behind
Location	Hub sites	Hub sites	Hub sites	Hub sites	Hub sites	Hub sites	Hub sites	Hub sites
Providers	District teachers	District teachers	District teachers	District teachers	District teachers	District teachers	District teachers	District teachers
Delivery	In person	In person	In person	In person	In person	In person	In person	In person
Intended Frequency	5 days a week	4 days a week	4 days a week	4 days a week	5 days a week	3 to 4 days a week	5 days a week	4 days a week
Intended Dosage	4 weeks, 19 days	4 weeks, 15 days	4 weeks, 15 days	5 weeks, 20 days	3 weeks, 15 days	12 to 18 days	4 weeks, 20 days	6 weeks, 17 days
Academic Time per Day, per Subject	90 minutes	90 minutes	45 to 90 minutes	90 to 100 minutes	90 minutes	60 to 120 minutes	60 minutes	120 minutes
Operating Hours	Extended Hours	Regular School Hours	Half School Day	Half School Day	Half School Day	Half School Day	Half School Day	Full School Day
Other Programming		Tutoring	Virtual summer program	Tutoring		Tutoring		

Notes: Based on interview notes with school district leaders conducted between Fall 2022 and early Spring 2023.

Table 3. Summer Programs: Participation and Dosage

District	Sample Rising Grade Level	Sample Size	% Treated			Intended Program Length (in Days)	# of Days Attended (Average)	% Total Days Attended[3]	Total Hours of Instruction per Student	
			Overall	Hub Sites[2]	Not Hub				Math	Reading
District 1	G3-8	11,841	22.6%	28.1%	20.3%	19	12.5	65.7%	18.7	18.7
District 2	G3-8	39,248	15.0%	19.8%	11.3%	15	10.8	72.1%	16.2	16.2
District 3	G3-8	5,924	46.2%[1]	NA[4]	NA[4]	15	NA	NA	NA	NA
District 4	G3-8	14,689	15.1%	17.7%	14.6%	20	11.6	58.0%	23.2	23.2
District 5	G1-8	1,804	13.6%	17.4%	14.8%	15	9.9	66.1%	14.9	14.9
District 6	G1-8 (math) G4-8 (reading)	33,504	12.1%	13.4%	10.5%	12-18	10.9	65.2%	20.2	21.0
District 7	G4-8	13,341	8.4%	18.8%	7.0%	20	14.5	72.5%	14.5	14.5
District 8	G6-8	9,370	4.8%	5.9%	4.5%	17	13.6	80.0%	27.2	27.2

Notes:
1. For District 3, "% Treated" is based on the % of students who signed up for summer school, which included students who never showed up. The estimate includes an upward bias.
2. "% Treated (Hub Sites)" shows the participation rate among students who were enrolled at a school that served as a summer school site.
3. "% Total Days Attended" is based on students who attended any summer school and given by dividing intended program length (in days) by the average number of days attended.
4. Hub site information was not available for District 3.

Table 4. Effects of Attending Summer Programs on Fall MAP Test, Math

District (Rising Grades)	Estimate: Any Participation (SE)	Estimate: Hourly (SE)	Avg Dosage in Days	Avg Dosage in Hours	Expected Effect from Research	% Treated	N
Overall	0.0240** (0.0078)	0.0014** (0.0005)	12.13	17.20	0.0378	10.70%	122,128
Overall (omitting D3 & 8)	0.0321** (0.0050)	0.0018** (0.0003)	11.57	17.81	0.0357	9.92%	108,408
District 1 (3-8)	0.0435** (0.0105)	0.0029** (0.0008)	12.52	18.78	0.0347	23.03%	11,220
District 2 (3-8)	0.0267** (0.0081)	0.0019** (0.0005)	10.84	15.42	0.0271	8.19%	37,361
District 3 (3-7)	0.0138 (0.0202)	0.0012 (0.0018)	15.00	11.25	0.0207	42.92%	4,350
District 4 (3-7)	0.0384** (0.0147)	0.0014** (0.0006)	11.85	23.89	0.0440	9.57%	12,074
District 5 (1-8)	0.0457 (0.0310)	0.0040* (0.0020)	9.91	14.86	0.0275	13.75%	1,767
District 6 (1-8)	0.0362** (0.0115)	0.00167** (0.0005)	10.23	18.68	0.0345	7.94%	32,645
District 7 (4-8)	-0.0003 (0.0185)	0.0002 (0.0013)	14.50	14.50	0.0268	8.40%	13,341
District 8 (6-8)	-0.0189 (0.0158)	-0.0008 (0.0006)	13.60	27.20	0.0503	4.78%	9,370

Notes: *p<0.05, **p<0.01
1. "Overall" refers to the meta-analytic estimates of the eight coefficients. The second overall estimate excludes Districts 3 and District 8.
2. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
3. Spring 2022 MAP testing was limited in District 7 and therefore we used non-missing Winter 2022 MAP and state standardized test to define the sample.

Table 5. Effects of Attending Summer Programs on Fall MAP Test, Reading

District (Rising Grades)	Estimate: Any Participation (SE)	Estimate: Hourly (SE)	Avg Dosage in Days	Avg Dosage in Hours	Expected Effect from Research	% Treated	N
Overall	-0.0032 (0.0079)	0.0000 (0.0004)	12.53	16.89	0.0314	11.24%	103,603
Overall (omitting D3 & 8)	0.0027 (0.0082)	0.0003 (0.0004)	11.89	17.69	0.0313	10.19%	89,054
District 1 (3-8)	-0.0032 (0.0118)	0.0002 (0.0009)	12.51	18.765	0.0309	22.39%	11,110
District 2 (3-8)	-0.0197 (0.0132)	-0.0004 (0.0008)	11.61	14.88	0.0245	8.57%	28,106
District 3 (3-8)	-0.029 (0.0172)	-0.0026 (0.0015)	15.00	11.25	0.0185	41.00%	5,175
District 4 (3-8)	0.0068 (0.0144)	0.0000 (0.0006)	11.65	23.23	0.0381	9.56%	14,418
District 5 (1-8)	0.0045 (0.0348)	0.0006 (0.0024)	9.95	14.93	0.0246	13.70%	1,744
District 6 (4-8)	0.0287* (0.0135)	0.0013* (0.0006)	9.85	18.18	0.0300	7.08%	20,418
District 7 (4-8)	0.0019 (0.0233)	-0.0007 (0.0011)	14.50	14.50	0.0239	8.40%	13,258
District 8 (6-8)	-0.0330 (0.0324)	-0.0009 (0.0011)	13.60	27.20	0.0448	4.78%	9,374

Notes: *p<0.05, **p<0.01
1. "Overall" refers to the meta-analytic estimates of the eight coefficients. The second overall estimate excludes Districts 3 and District 8.
2. For District 3, we estimated the effects of signing up for the summer program. The listed dosages are the maximum number of days and hours students could attend.
3. In District 6, rising grades 1-3 are excluded from the sample because of low reading testing rates in those grades.
4. Spring 2022 MAP testing was limited in District 7 and therefore we used non-missing Winter 2022 MAP and state standardized test to define the sample.

Table 6. Effects of Summer Programs and Tutoring on Fall MAP Test, Math

District (Rising Grades)	Estimate: Any Participation (SE)	Estimate: Hourly (SE)	Avg Tutoring Dosage in Hours	Avg Summer Program Dosage in Hours	Expected Effect from Research	% Treated	N
Overall	0.0409** (0.0092)	0.0017** (0.0004)	18.73	6.88	0.0817	4.99%	82,080
District 2 (3-8)	0.0384** (0.0115)	0.0017** (0.0005)	22.01	3.92	0.0878	6.63%	37,361
District 4 (3-7)	0.0470* (0.0210)	0.0021* (0.0007)	23.46	0.00	0.0826	6.84%	12,074
District 6 (1-8)	0.0439 (0.0235)	0.0009 (0.0008)	3.51	23.33	0.0555	2.42%	32,645

Notes: *p < 0.05; **p < 0.01.

1. SE is standard error.

Table 7. Effects of Summer Programs and Tutoring on Fall MAP Test, Reading

District (Rising Grades)	Estimate: Any Participation (SE)	Estimate: Hourly (SE)	Avg Tutoring Dosage in Hours	Avg Summer Program Dosage in Hours	Expected Effect from Research	% Treated	N
Overall	-0.0086 (0.0151)	0.0000 (0.0000)	21.03	5.42	0.0714	5.09%	62,942
District 2 (3-8)	-0.0180 (0.0210)	-0.0003 (0.0007)	24.00	4.10	0.0788	6.00%	28,106
District 4 (3-7)	-0.0197 (0.0302)	-0.0004 (0.0010)	23.41	0.00	0.0759	7.59%	14,418
District 6 (4-8)	0.0235 (0.0309)	0.0012 (0.0010)	3.12	24.64	0.0507	2.08%	20,418

Notes: *p < 0.05, **p < 0.01.

1. SE is standard error.

Table 8. Effects of Summer Programs on Fall MAP Test by Grade Bands, Math

District (Rising Grades)	Any Participation (SE)	Average Dosage in Hours	% Treated	N
District 1 (3-5)	0.0539**	19.09	27.37%	6,368
	(0.0149)			
District 2 (3-5)	0.0390**	15.65	9.00%	19,265
	(0.0110)			
District 3 (3-5)	0.0188	11.25	43.44%	2,958
	(0.0270)			
District 4 (3-5)	0.0590**	24.79	10.00%	7,354
	(0.0210)			
District 5 (1-5)	0.0457	17.02	16.54%	1,028
	(0.0403)			
District 6 (1-5)	0.0493**	20.97	8.92%	20,525
	(0.0147)			
District 7 (4-5)	0.0047	14.36	11.86%	5,470
	(0.0254)			
District 1 (6-8)	0.0247*	18.14	17.33%	4,852
	(0.0106)			
District 2 (6-8)	0.0100	15.13	7.00%	18,096
	(0.0120)			
District 3 (6-8)	0.0038	11.25	41.81%	1,392
	(0.0278)			
District 4 (6-8)	-0.0040	22.20	8.00%	4,720
	(0.0150)			
District 5 (6-8)	0.0457	9.84	9.88%	739
	(0.0512)			
District 6 (6-8)	0.0069	13.16	6.29%	12,120
	(0.0171)			
District 7 (6-8)	-0.0059	14.35	6.09%	7,871
	(0.0274)			
District 8 (6-8)	-0.0189	27.20	4.78%	9,370
	(0.0158)			

Notes: *p < 0.05; **p < 0.01.

1. SE is standard error.

Table 9. Effects of Summer Programs on Fall MAP Test by Grade Bands, Reading

District (Rising Grades)	Any Participation (SE)	Average Dosage in Hours	% Treated	N
District 1 (3-5)	-0.0005 (0.0143)	19.00	26.69%	6,318
District 2 (3-5)	-0.0120 (0.0160)	14.72	9.00%	12,015
District 3 (3-5)	-0.0275 (0.0186)	11.25	43.45%	2,939
District 4 (3-5)	-0.0070 (0.0210)	25.05	12.00%	7,315
District 5 (1-5)	0.0076 (0.0428)	17.08	16.88%	995
District 6 (4-5)	0.0239 (0.0213)	20.29	8.43%	8,457
District 7 (4-5)	0.0225 (0.0270)	14.31	11.74%	5,477
District 1 (6-8)	-0.0076 (0.0210)	18.25	16.74%	4,792
District 2 (6-8)	-0.0270 (0.0200)	15.00	8.00%	16,091
District 3 (6-8)	-0.0310 (0.0322)	11.25	0.3779	2,236
District 4 (6-8)	0.0290 (0.0170)	20.19	7.00%	7,103
District 5 (6-8)	-0.0020 (0.0677)	9.85	9.48%	749
District 6 (6-8)	0.0325 (0.0170)	16.12	6.12%	11,961
District 7 (6-8)	-0.0218 (0.0394)	14.30	6.00%	7,781
District 8 (6-8)	-0.0330 (0.0324)	27.20	4.78%	9,374

Notes: $^{*}p < 0.05$; $^{**}p < 0.01$.

1. SE is standard error.

Appendix Table A1. Sample Characteristics by Treatment Status

	District 1		District 2		District 3[5]		District 4		District 5		District 6		District 7		District 8	
	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated	Treated	Not Treated
Special Education[1]	16.0%	14.0%	25.9%	16.9%	16.1%	8.2%	39.7%	28.1%	19.2%	17.9%	22.2%	10.3%	35.9%	18.0%	50.9%	14.1%
ELL	34.0%	36.2%	57.4%	52.6%	61.0%	21.7%	45.7%	29.2%	20.4%	11.4%	24.1%	12.2%	20.4%	11.9%	26.6%	9.8%
FRPL[2]	72.3%	77.2%	94.4%	87.0%	NA	NA	73.9%	52.1%	42.9%	34.3%	NA	NA	NA	NA	NA	NA
Male	48.7%	51.3%	50.3%	50.4%	52.8%	51.4%	51.9%	51.3%	53.5%	47.7%	52.9%	50.7%	53.0%	50.9%	57.8%	49.4%
Race:																
Black	24.2%	20.3%	18.5%	14.5%	25.7%	24.6%	29.0%	18.6%	13.5%	10.1%	52.6%	39.9%	19.7%	6.0%	67.4%	27.8%
Hispanic	37.0%	41.2%	78.0%	75.1%	52.5%	26.1%	46.4%	35.4%	0.4%	0.2%	23.8%	17.9%	29.1%	15.0%	25.9%	11.7%
Asian	1.7%	1.8%	0.7%	1.0%	7.6%	6.8%	8.6%	6.6%	13.1%	8.0%	5.0%	7.5%	4.5%	5.9%	0.0%	1.8%
Other	15.9%	16.3%	2.1%	3.3%	1.4%	3.8%	2.1%	3.4%	32.2%	22.8%	5.4%	5.6%	15.8%	13.0%	1.3%	3.8%
White	21.2%	20.4%	2.1%	7.7%	12.9%	38.7%	13.9%	36.0%	40.8%	58.8%	13.2%	29.2%	30.9%	60.2%	5.4%	54.9%
Priority for summer[3]	NA	NA	72.9%	29.6%	64.9%	3.1%	51.5%	24.1%	NA	NA	92.2%	50.1%	100.0%	34.8%	NA	NA
Spring 2022 MAP test[4]																
Math (RIT points)	187.6	191.8	192.9	203.9	184.8	207.9	193.6	207.3	198.1	206.4	177.3	197.5	190.4	212.4	197.0	226.6
Math (normed)	-1.02	-1.00	-0.78	-0.15	-1.28	-0.06	-0.64	0.11	-0.05	0.10	-1.30	-0.25	-1.18	0.02	-1.52	0.16
Reading (RIT points)	191.5	195.2	199.5	208.7	188.6	209.0	199.9	213.1	191.5	200.8	179.5	200.1	191.3	213.5	196.3	224.2
Reading (normed)	-0.87	-0.82	-0.81	-0.25	-1.14	0.19	-0.65	0.12	-0.12	0.08	-1.29	-0.20	-0.80	0.38	-1.53	0.21
Fall 2022 MAP test																
Math (RIT points)	187.9	192.3	193.0	203.1	188.0	210.8	195.2	208.1	196.3	204.8	181.7	200.2	194.3	215.7	190.4	218.3
Math (normed)	-0.93	-0.92	-0.67	-0.11	-0.76	0.28	-0.54	0.15	0.77	0.79	-0.93	0.01	-1.01	0.17	-1.36	0.24
Reading (RIT points)	191.1	194.6	199.4	207.5	191.4	210.3	197.8	209.8	191.4	200.6	189.2	207.4	195.1	216.2	190.1	217.7
Reading (normed)	-0.79	-0.73	-0.78	-0.21	-0.87	0.38	-0.51	0.18	0.63	0.71	-0.95	0.00	-0.64	0.45	-1.45	0.21
N	2,671	9,170	5,881	33,367	2,734	3,190	2,222	12,467	245	1,559	4,037	29,467	1,162	12,587	448	8,926

Notes:

1. "Special Education" includes Plan 504 students.
2. FRPL refers to "free or reduced priced lunch." Student level FRPL status data was not available in District 3, 6, 7, and 8.
3. "Priority for summer" refers to students who were prioritized to receive summer programming based on criteria that varied by district. District 1: no priority students. District 2: students who scored below grade-by-subject threshold on state tests. District 3: students identified as priority based on district's own prioritization matrix. District 4: students who scored below grade-by-subject threshold on state tests. District 5: no priority students. District 6: low-scoring students. District 7: low-performing and historically underserved students. District 8: Students identified as academically behind. For District 8, we do not report the statistics since the criteria for priority were based on teacher discretion.
4. For District 7, scores are from Winter 2022 rather than Spring 2022 due to limited availability of Spring tests in the district.
5. For District 3, the definition of "Treated" is signing up for summer program, not attending summer program.

Appendix Table A2. District 1: Subgroup Analysis Results

	Math (Rising Grades 3-8)				Reading (Rising Grades 3-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.0435** (0.0105)	18.78	23.03%	11,220	-0.0032 (0.0118)	18.77	22.39%	11,110
Below 50th Percentile	0.0371* (0.0166)	18.24	23.03%	5,502	0.0169 (0.0184)	18.51	22.69%	5,377
Above 50th Percentile	0.0366* (0.0147)	19.31	23.03%	5,718	-0.0345 (0.0174)	19.00	22.12%	5,733
Not FRPL	0.0583* (0.0247)	19.62	26.68%	2,736	-0.0080 (0.024)	19.61	25.55%	2,658
FRPL	0.0365** (0.0120)	18.45	21.85%	8,484	-0.0017 (0.0134)	18.44	21.40%	8,452
Not Special Education	0.0292** (0.0107)	18.65	22.67%	9,591	-0.0056 (0.0125)	18.66	21.99%	9,523
Special Education	0.1056** (0.0350)	19.55	25.17%	1,629	-0.0057 (0.0389)	19.29	24.83%	1,587
Not ELL	0.0460** (0.0138)	18.66	23.61%	7,254	-0.0138 (0.0158)	18.59	22.92%	7,196
ELL	0.0337 (0.0203)	19.02	21.96%	3,966	0.0213 (0.0199)	19.09	21.44%	3,914
Female	0.0396** (0.0000)	18.67	24.00%	5,516	-0.0161 (0.0159)	18.63	23.24%	5,494
Male	0.0448** (0.0000)	18.90	22.09%	5,704	-0.0017 (0.0172)	18.90	21.56%	5,616
Black	0.0325 (0.0225)	17.52	26.11%	2,375	-0.0013 (0.0274)	17.34	25.64%	2,360
Hispanic	0.0230 (0.0177)	18.84	21.32%	4,488	0.0105 (0.0184)	18.86	20.60%	4,418
Asian, Multi-racial, NHPI, Native	0.0495 (0.0298)	18.74	22.31%	2,013	-0.0115 (0.0333)	18.67	22.16%	2,013
White	0.0675* (0.0282)	20.13	23.81%	2,344	0.0008 (0.0336)	20.29	22.73%	2,319

Notes: *p<0.05, **p<0.01
Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.

Appendix Table A3. District 2: Subgroup Analysis Results

	Math (Rising Grades 3-8)				Reading (Rising Grades 3-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.0267** (0.0081)	15.42	8.2%	37,361	-0.0197 (0.0132)	14.88	8.6%	28,106
Below 50th Percentile	0.0402** (0.0117)	15.55	9.9%	18,803	-0.0236 (0.0201)	15.13	10.7%	14,016
Above 50th Percentile	0.0066 (0.0128)	15.23	6.5%	18,558	-0.0144 (0.0165)	14.47	6.5%	14,090
Not FRPL	0.0727* (0.0317)	13.46	5.4%	5,294	-0.0621 (0.0349)	13.93	6.0%	4,833
FRPL	0.0225* (0.0088)	15.63	8.6%	32,067	-0.0155 (0.0147)	15.01	9.1%	23,273
Not Special Education	0.0329** (0.0084)	15.53	8.0%	30,709	-0.0157 (0.0154)	14.91	8.3%	22,473
Special Education	0.0084 (0.0235)	15	9.1%	6,652	-0.0235 (0.0287)	14.75	9.5%	5,633
Not ELL	0.0421** (0.0133)	13.94	7.3%	17,426	-0.0091 (0.0157)	14.39	8.6%	16,892
ELL	0.0195 (0.0108)	16.49	8.9%	19,935	-0.0322 (0.0237)	15.61	8.6%	11,214
Female	0.0283* (0.0112)	15.44	8.0%	18,540	-0.0227 (0.0173)	15.25	9.1%	13,849
Male	0.0292* (0.0117)	15.41	8.4%	18,821	-0.0200 (0.0191)	14.47	8.0%	14,257
Black	0.0263 (0.0246)	13.65	9.6%	5,440	-0.0014 (0.0265)	14.04	11.2%	5,404
Hispanic	0.0225* (0.0092)	15.91	8.5%	28,157	-0.0339* (0.0162)	15.22	8.7%	19,065
Asian, Multi-racial, NHPI, Native	NA	14.01	7.0%	1,177	NA	14.75	6.8%	1,144
White	NA	13.63	2.7%	2,587	NA	14.14	3.1%	2,493

Notes: *p<0.05, **p<0.01
Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.

Appendix Table A4. District 3: Subgroup Analysis Results

	Math (Rising Grades 3-7)				Reading (Rising Grades 3-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.0138 (0.0202)	11.25	42.92%	4,350	-0.029 (0.0026)	11.25	41.00%	5,175
Below 50th Percentile	-0.0229 (0.0261)	11.25	70.66%	1,994	-0.0331 (0.0258)	11.25	68.39%	2,392
Above 50th Percentile	0.0266 (0.0252)	11.25	19.44%	2,356	-0.026 (0.0237)	11.25	17.46%	2,783
>50% FRPL Schools (2019)	0.0200 (0.0252)	11.25	49.36%	2,960	-0.0408 (0.0196)	11.25	48.02%	3,353
<50% FRPL Schools (2019)	-0.0022 (0.0431)	11.25	29.21%	1,390	0.0396 (0.0280)	11.25	28.12%	1,821
Not Special Education	0.0090 (0.0231)	11.25	42.42%	4,029	-0.0321 (0.0175)	11.25	40.40%	4,770
Special Education	NA	11.25	49.22%	321	NA	11.25	48.15%	405
Not ELL	0.0013 (0.0282)	11.25	26.57%	2,589	-0.0663 (0.0259)	11.25	25.40%	3,236
ELL	0.0118 (-0.0322)	11.25	66.93%	1,760	0.0089 (0.0295)	11.25	67.04%	1,939
Female	0.0400 (0.0321)	11.25	42.71%	2,091	-0.0247 (0.0239)	11.25	40.45%	2,492
Male	-0.0068 (0.0241)	11.25	43.12%	2,259	-0.0495 (0.0239)	11.25	41.52%	2,683
Black	0.0323 (0.0372)	11.25	46.27%	1,113	0.0109 (0.0430)	11.25	44.16%	1,309
Hispanic	0.0272 (0.0260)	11.25	58.81%	1,515	-0.0182 (0.0354)	11.25	57.42%	1,846
Asian, Multi-racial, NHPI, Native	NA	11.25	42.13%	470	NA	11.25	38.20%	534
White	0.0342 (0.0428)	11.25	21.01%	1,252	-0.1185 (0.0537)	11.25	18.84%	1,486

Notes: *p<0.05, **p<0.01
1. Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.
2. For District 3, we estimated the effects of signing up for summer program. The listed dosages are the maximum number of days and hours students could attend.

Appendix Table A5. District 4: Subgroup Analysis Results

Subgroups	Math (Rising Grades 3-7) Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Reading (Rising Grades 3-8) Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.0384**	23.89	9.6%	12,074	0.0068	23.23	9.6%	14,418
	(0.0147)				(0.0144)			
Below 50th Percentile	0.0412*	23.99	13.5%	5,880	0.0023	24.54	13.3%	7,058
	(0.0182)				(0.0196)			
Above 50th Percentile	0.0099	23.68	5.9%	6,194	0.0050	20.46	6.0%	7,360
	(0.0206)				(0.0208)			
Not FRPL	0.0377	23.04	5.7%	5,423	0.0416	21.17	6.7%	6,484
	(0.0250)				(0.0234)			
FRPL	0.0399*	24.2	12.8%	6,651	-0.0128	24.17	11.9%	7,934
	(0.0181)				(0.0175)			
Not Special Education	0.0496**	23.6	9.2%	8,437	-0.0068	22.79	9.2%	10,111
	(0.0160)				(0.0156)			
Special Education	0.0257	24.49	10.4%	3,637	0.0436	24.14	10.5%	4,307
	(0.0296)				(0.0284)			
Not ELL	0.0271	23.6	7.6%	8,176	0.0293	22.55	8.5%	9,897
	(0.0173)				(0.0182)			
ELL	0.0555*	24.23	13.7%	3,898	-0.0245	24.31	11.9%	4,521
	(0.0235)				(0.0214)			
Female	0.0063	23.87	9.5%	5,888	0.0137	23.42	10.1%	7,014
	(0.0195)				(0.0207)			
Male	0.0773**	23.91	9.7%	6,186	0.0039	23.04	9.1%	7,404
	(0.0210)				(0.0190)			
Black	0.0322	23.83	13.8%	2,430	0.0067	23.52	14.0%	2,907
	(0.0297)				(0.0337)			
Hispanic	0.0565**	24.31	11.9%	4,401	0.0259	24.24	11.0%	5,300
	(0.0214)				(0.0220)			
Asian, Multi-racial, NHPI, Native	0.0836	22.91	11.2%	1,234	-0.0864	22.92	10.8%	1,466
	(0.0494)				(0.0468)			
White	-0.0300	23.48	3.9%	4,009	0.0012	20.41	4.9%	4,745
	(0.0340)				(0.0296)			

Notes: *p<0.05, **p<0.01
Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.

Appendix Table A6. District 6: Subgroup Analysis Results

	Math (Rising Grades1-8)				Reading (Rising Grades 4-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.0362** (0.0115)	18.68	7.94%	32,645	0.0287* (0.0135)	18.18	7.08%	20,418
Below 50th Percentile	0.0332** (0.0126)	19.36	16.32%	17,528	0.0283 (0.0149)	18.96	14.39%	10,886
Above 50th Percentile	0.0347 (0.0336)	21.43	2.02%	15,117	0.0068 (0.0363)	17.70	1.40%	9,532
Not FRPL	NA	NA	NA	NA	NA	NA	NA	NA
FRPL	NA	NA	NA	NA	NA	NA	NA	NA
Not Special Education	0.0327** (0.0126)	19.26	8.52%	28,845	0.0321* (0.0154)	18.42	7.23%	17,949
Special Education	0.0483 (0.0339)	20.61	18.61%	3,800	-0.0014 (0.0344)	20.29	16.24%	2,469
Not ELL	0.0359** (0.0125)	19.30	8.40%	28,172	0.0266 (0.0155)	18.67	7.28%	17,648
ELL	0.0332 (0.0266)	20.32	17.84%	4,473	0.0577 (0.0334)	19.44	14.98%	2,770
Female	0.0256 (0.0164)	19.54	9.27%	15,983	0.0263 (0.0198)	19.16	8.13%	9,974
Male	0.0389* (0.0162)	19.57	10.10%	16,662	0.0323 (0.0191)	18.59	8.50%	10,444
Black	0.0423** (0.0154)	18.63	12.34%	13,422	0.0224 (0.0187)	18.49	11.38%	8,358
Hispanic	0.0070 (0.0244)	19.62	13.02%	6,051	0.0464 (0.0290)	18.97	10.93%	3,842
Asian, Multi-racial, NHPI, Native	0.0416 (0.0393)	22.07	7.54%	4,069	0.0216 (0.0512)	20.19	5.54%	2,565
White	0.0484 (0.0348)	21.31	4.43%	8,987	0.0003 (0.0490)	19.51	3.22%	5,594

Notes: *p<0.05, **p<0.01
Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.

Appendix Table A7. District 7: Subgroup Analysis Results

	Math (Rising Grades 4-8)				Reading (Rising Grades 4-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	-0.0003 (0.0185)	14.5	8.40%	13,341	0.0019 (0.0233)	14.50	8.40%	13,258
Below 50th Percentile	-0.0021 (0.0196)	14.35	16.54%	6,205	0.0033 (0.0279)	14.19	19.72%	4,290
Above 50th Percentile	-0.0613 (0.0511)	14.50	1.44%	7,136	0.0024 (0.037)	14.70	2.94%	8,968
Not FRPL	NA	NA	NA	NA	NA	NA	NA	NA
FRPL	NA	NA	NA	NA	NA	NA	NA	NA
Not Special Education	-0.0013 (0.0226)	14.49	6.74%	10,750	-0.0340 (0.0285)	14.44	6.66%	10,709
Special Education	0.0389 (0.0380)	14.12	15.59%	2,591	0.0873 (0.0463)	14.09	15.57%	2,549
Not ELL	-0.0029 (0.0402)	14.29	7.69%	1,672	0.019 (0.0222)	14.23	7.60%	11,619
ELL	0.0015 (0.0211)	14.62	13.88%	11,669	-0.0503 (0.0532)	14.63	13.85%	1,639
Female	0.0274 (0.0236)	14.62	8.12%	6,529	-0.0212 (0.0298)	14.58	8.06%	6,499
Male	-0.0274 (0.0288)	14.13	8.79%	6,812	0.0106 (0.0339)	14.08	8.67%	6,759
Black	NA	14.12	23.35%	938	NA	13.96	23.37%	933
Hispanic	0.0126 (0.0314	14.05	15.23%	2,147	-0.0295 (0.0399)	14.03	15.38%	2,106
Asian, Multi-racial, NHPI, Native	0.0229 (0.0413)	14.29	9.15%	2,513	-0.0170 (0.0481)	14.25	8.98%	2,528
White	-0.0264 (0.0355)	14.84	4.56%	7,743	0.117*** (0.0349)	14.85	4.43%	7,691

Notes: *p<0.05, **p<0.01
1. Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.
2. District 7 did not have available data on FRPL status.

Appendix Table A8. District 8: Subgroup Analysis Results

	Math (Rising Grades6-8)				Reading (Rising Grades 6-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	-0.019	27.2	4.78%	9,266	-0.033	27.2	4.78%	9,242
	(0.0158)				(0.0324)			
Below 50th Percentile	-0.0238	27.22	10.30%	4,049	-0.0329	27.22	11.75%	3,610
	(0.0164)				(0.0332)			
Above 50th Percentile	NA	31.07	0.29%	5,217	NA	11.95	0.27%	5,632
Not FRPL	NA	NA	NA	NA	NA	NA	NA	NA
FRPL	NA	NA	NA	NA	NA	NA	NA	NA
Not Special Education	-0.0304	27.86	2.74%	1,462	-0.0416	27.68	2.75%	7,775
	(0.0252)				(0.0321)			
Special Education	0.0012	26.85	14.91%	7,804	-0.0117	26.77	15.34%	1,467
	(0.0195)				(0.0512)			
Not ELL	-0.0267	27.54	3.80%	8,297	-0.0201	27.43	3.90%	8,281
	(0.0195)				(0.0514)			
ELL	NA	26.86	12.07%	969	NA	26.62	12.07%	961
Female	0.0152	27.96	4.03%	4,516	-0.0329	27.56	4.14%	4,497
	(0.0205)				(0.0486)			
Male	-0.0481	26.91	5.56%	4,749	-0.0315	26.96	5.33%	4,744
	(0.0253)				(0.037)			
Black	-0.0171	27.94	10.57%	2,724	-0.0362	27.83	10.95%	2,713
	(0.0232)				(0.0314)			
Hispanic	0.0065	26.93	10.05%	1,134	0.0162	26.69	9.97%	1,133
	(0.0365)				(0.0466)			
Asian, Multi-racial, NHPI, Native	NA	15.59	1.20%	499	NA	15.59	1.12%	497
White	NA	23.67	0.49%	4,909	NA	23.21	0.47%	4,899

Notes: *p<0.05, **p<0.01
Only impact estimates for subgroups with N>1,000 and the number of treated observations >100 are reported.
Figure 1. Summer Programs: Effect of Treatment on Math Outcomes