Summer School as an Academic Recovery Strategy After COVID-19: Evidence from Summer 2023

Atticus Bolyard
Michael DeArmond
Elise Dizon-Ross
Dan Goldhaber
Jazmin Isaacs
Thomas J. Kane
Micaela Keating
Jamie Klinenberg
Anna McDonald
Andrew McEachin
Emily Morton
Atsuko Muroga
Alejandra Salazar
Douglas O. Staiger

August 2025

WORKING PAPER No. 325-0825

CALDER
National Center for Analysis of Longitudinal Data in Education Research

AIR®

Summer School as an Academic Recovery Strategy After COVID-19: Evidence from Summer 2023

Atticus Bolyard
Harvard University

Thomas J. Kane
Harvard University

Micaela Keating
Harvard University

Atsuko Muroga
Harvard University

Michael DeArmond
American Institutes for Research / CALDER

Elise Dizon-Ross
American Institutes for Research / CALDER

Dan Goldhaber
American Institutes for Research / CALDER

Jamie Klinenberg
American Institutes for Research / CALDER

Anna McDonald
American Institutes for Research / CALDER

Alejandra Salazar
American Institutes for Research / CALDER

Jazmin Isaacs
NWEA

Emily Morton
NWEA

Andrew McEachin
Educational Testing Service

Douglas O. Staiger
Dartmouth College

Contents

Acknowledgments

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 Lisa Sanbonmatsu for her valuable contributions to this work.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Summer School as an Academic Recovery Strategy After COVID-19: Evidence from Summer 2023

Atticus Bolyard, Michael DeArmond, Elise Dizon-Ross, Dan Goldhaber, Jazmin Isaacs, Thomas J. Kane, Micaela Keating, Jamie Klinenberg, Anna McDonald, Andrew McEachin, Emily Morton, Atsuko Muroga, Alejandra Salazar, & Douglas O. Staiger

CALDER Working Paper No. 325-0825
August 2025

Abstract

Many school districts have turned to summer school as a strategy for addressing persistent declines in student achievement after COVID-19. We examine the impact of 2023 summer programs on academic recovery across 400,000 students in eight large U.S. districts. Using value-added models, we find modest but significant improvements in student math achievement (average effect = 0.024 SD), and no significant improvements in reading. Effects were similar to 2022, despite small declines in participation and dosage. In four districts, we also examine differences in summer 2023 impacts for students who attended the districts' summer school the prior year and those who did not. Students attending for a second consecutive summer generally benefited as much or more than first-time attendees.

1. Introduction

Despite public perceptions that students have recovered or will recover from the pandemic's impact on student achievement (Polikoff et al., 2023), multiple assessments show only slight improvements in math and stagnant or declining reading scores as of spring 2024 (Dewey et al., 2024b; Fahle et al., 2024; Lewis & Kuhfeld, 2023; 2024; NCES, 2025). Low-performing and minoritized students have been disproportionately affected, widening existing achievement gaps. Federal investments in academic recovery following the pandemic (i.e., the Elementary and Secondary School Emergency Relief Fund, ESSER) positively impacted student achievement (Dewey et al., 2024a; Goldhaber & Falken, 2025); nevertheless, at least two-thirds of the initial decline in math and nearly all of the initial decline in reading remain unaddressed (Dewey et al., 2024b; NCES, 2025). Interestingly, most academic recovery since spring 2021 has occurred during summer periods, driven by reduced summer slide (Lewis & Kuhfeld, 2022, 2023). As ESSER funding ends and promising interventions like tutoring and double-dose courses struggle to reach enough students to offset pandemic-related learning losses (Carbonari et al., 2024a, 2024b; Kraft et al., 2024a), districts urgently need information about pandemic-era programs that can improve achievement at scale. With that in mind, this paper examines a widespread program that holds promise as a scalable recovery initiative: summer school.

1.1 Impacts of Summer School Pre- and Post-Pandemic

Following the first full pandemic-impacted school year, many school districts launched or increased the scale of their summer programs in 2022 with ESSER funds (DiMarco & Jordan, 2022; Roche, 2023; Diliberti & Schwartz, 2022). This expansion was supported by pre-pandemic research showing summer school's effectiveness in improving math (Augustine et al., 2016; Lynch et al., 2023) and reading achievement (Kim & Quinn, 2013). However, more recent studies of pandemic-era summer school have produced mixed results, with most finding a

positive association between summer school participation and test score gains, particularly in math (e.g., Barry & Sass, 2022; Borman et al., 2024; [redacted]). These studies also find the programs were shorter and had fewer total instructional hours than recommended in pre-pandemic research (Schwartz et al., 2018). Perhaps unsurprisingly, the pandemic-era programs have had smaller average effects, enough to offset only a fraction of districts' achievement declines. For example, the 0.03 SD math gain associated with 2022 summer school participation across eight districts in [redacted] is equivalent to just 2% of the districts' declines in math (accounting for the fact that only 13% of the districts' students participated).

While the small impact of pandemic-era summer school may seem underwhelming, it is important to consider that summer school served a much larger share of students than many school year recovery programs. In recent studies of academic recovery programs delivered during 2021-22 and 2022-23, for example, no intervention that had significant positive effects reached more than 1% of a district's K-8 students (Carbonari et al., 2024a, 2024b). Moreover, the robust evidence base on tutoring, one of the most common school year recovery interventions, primarily comes from small randomized control trials—few involving more than 400 students—and program effects drastically decline as scale increases (Kraft et al., 2024b). Whereas the widespread enthusiasm for high-intensity tutoring as a pandemic recovery strategy was based on impressive results in limited settings (e.g., Guryan et al., 2023; Jacobson, 2024), summer school arguably has more evidence for effectiveness *at scale* and the potential to meaningfully contribute to district- and nation-wide academic recovery.

An important question is whether students benefit from multiple years of summer school, or if repeated participation yields diminishing returns. Despite the relevance of this question for policymakers and practitioners, evidence on the impact of attending multiple years of summer

school is scarce. Augustine et al. (2016) examined effects of being invited to two consecutive years of summer school across five urban districts. They found positive effects of the initial invitation on fall math (but not reading) scores after the first summer (+0.08 SD), but no impacts after the second. While 79% of invited students attended the first summer, only 52% participated both years. This drop in participation diluted the treatment rate for their intent-to-treat analysis, making it harder to detect any causal effect of two consecutive summers. When analyzing only the students who attended the second summer relative to the control group students, Augustine et al. (2016) find suggestive, correlational evidence that the second summer participants' scores improved in both math and language arts. Taken together, these findings suggest each summer independently had a positive effect on student achievement. However, it remains unknown if students gained more from attending two summers than from participating in only the second.

This study examines summer school in two ways. First, we estimate achievement impacts of 2023 summer programs across eight large districts. Second, in four of the districts, we examine the extent to which students who participate in two consecutive years of summer school benefit from the second summer. With the end of ESSER in September 2024, these results provide critical information to districts and states making difficult decisions about how to allocate increasingly scarce resources to address persisting pandemic learning losses.

1.2 Research Questions

This study addresses two questions:

> RQ1. What was the effect of participating in 2023 summer school on student achievement, and how does it compare to the effect of 2022 summer school?
>
> RQ2. To what extent did the effect of attending 2023 summer school differ for students who also attended summer school the previous summer compared to those who did not?

2. Method

2.1 Data and Sample

This paper draws on data from eight large districts participating in the [redacted] project during summer 2023, six of which also participated in summer 2022 (see [redacted]). Collectively, the eight districts enroll over 400,000 students across seven states. As shown in Table 1, the districts serve higher percentages of Black and Hispanic students (69%) and students eligible for free and reduced-price lunch (76%) than national averages (44% and 54%, respectively). Though there was some variation, most [redacted] districts experienced similar or larger test score declines from 2019 to 2022 than the national average (Appendix Table A1). From 2022 to 2024, recovery varied by district and subject but generally lagged or matched national trends.

Data include (a) summer program design details from district leader interviews, (b) student-level eligibility and participation data, (c) NWEA MAP Growth test scores, and (d) district-level recovery estimates on state tests in math and ELA from the Education Recovery Scorecard (Reardon et al., 2025).

Districts use MAP Growth assessments to monitor elementary and secondary students' reading and math achievement gains in the fall, winter, and spring. Further details about MAP Growth's reliability and validity are available in the 2019 MAP Growth technical report (NWEA, 2019). We standardize scores by subject and grade using the NWEA 2020 MAP Growth norms (Thum & Kuhfeld, 2020), which are based on a nationally representative pre-pandemic sample (2015–16 to 2017–18). We define grade level as a student's expected rising grade in fall 2023 based on their spring 2023 grade.

For RQ1, the analytic sample includes students entering grades 1-8 in fall 2023 who were eligible to participate in summer school (based on their grade level and school) and who have

MAP Growth scores in both spring and fall 2023 in reading or math. Appendix Table A2 shows demographics and test scores for participants and non-participants in each district. Participants were more likely to qualify for special education services, be English Learners, and be Black or Hispanic. In all districts, participants had lower (or equal) baseline MAP Growth scores than non-participants.

For RQ2, we focus on the four districts (3, 4, 7, and 8) with at least 100 students who participated in both 2022 and 2023 summer school. The analytic sample includes students entering grades 1-8 in fall 2023 who (a) have MAP Growth scores in spring 2022 and fall 2023 in either subject and (b) were in grades that were eligible for summer school in both years. For the purpose of this study, we label students who attended in both years "repeaters" and those who only attended in 2023 "first-timers." First-timers may have previously attended summer school in earlier years or other districts, but we do not have this information in our data. The repeater and first-timer samples are described in Appendix Table A3. In three of the four districts, students who did not attend either summer had the highest baseline and outcome MAP Growth scores (normed) in both subjects. Repeaters typically had the lowest baseline and outcome scores—with the exception of District 7, where they had the highest scores.

2.2 Empirical Strategy

For RQ1, 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:

$$MAP_{subj,ijt} = \beta_0 + \beta_1 SS23_i + \tau \boldsymbol{X}_{jt} \times Grade_{jt} + \varphi Sch\_Grade_{ij,t-1} + \varepsilon_{ijt} \qquad (1)$$

where $MAP_{subj,ijt}$ denotes the standardized math or reading MAP Growth score for student *i* in school *j* in term *t,* where *t* is fall 2023. Our main treatment variable is the $SS23_i$ term, a binary

indicator equal to 1 if student *i* attended at least one day of summer school in 2023 (we describe actual participation rates in more detail below). The vector $\mathbf{X}_i$ includes students' prior achievement, demographics, and missing data indicators. Specifically, we include student *i*'s prior same-subject MAP scores from spring 2023 (as a cubic polynomial), winter 2023, fall 2022, and their opposite-subject MAP score from spring 2023. We interact $\mathbf{X}_i$ with student *i*'s grade and include school-by-grade fixed effects.

The coefficient of interest is β_1, which represents the effect of participating in summer school on fall MAP Growth scores. For β_1 to be interpreted as a causal effect, it must be assumed that assignment to summer school is as good as random, conditional on the covariates included in our model. Our approach aligns with value-added methodologies that have been validated for generating causal estimates of teacher and school effects, such as incorporating multiple prior test scores in our covariates (e.g., Abdulkadiroglu et al., 2011; Bacher-Hicks et al., 2019; Chetty et al., 2014; Deming, 2014). Still, because our study relies on a selection-on-observables design, there is a risk of selection bias if participation is influenced by unmeasured factors that also affect future achievement. The direction of this bias is uncertain. For example, if more motivated or well-resourced (e.g., access to transportation) students are more likely to attend summer school, our estimates may be positively biased; if struggling or lower-resourced students were more likely to enroll, the bias could be negative.

For RQ2, we add an interaction term to Equation #1, and a control for attending summer school in 2022, to estimate the differential effect of summer school in 2023 on students who also attended summer school in 2022:

$$MAP_{subj,ijt} = \beta_0 + \beta_1 SS22_i + \beta_2 SS23_i + \beta_3 [SS22_i \times SS23_i] + \tau \boldsymbol{X}_{jt} \times Grade_{jt} + \varphi Sch_{Grade_{ij,t-1}} + \varepsilon_{ijt} \quad (2)$$

The coefficient on $SS22_i$ (β_1) represents the association between having participated in summer programming in 2022 but *not* 2023 and fall 2023 MAP scores. The coefficient on $SS23_i$ (β_2) represents the effect of participating only in 2023 (first-timers). The coefficient on the interaction term (β_3) is the extent to which repeaters benefited more or less from participating in summer 2023 programming than first-timers. A positive β_3 would indicate that, on average, repeaters benefited more from attending summer school in 2023 than first-timers.

3. Results

3.1 Program Designs

The designs of the districts' primary academic summer programs are described in Appendix Table A4. Six of eight districts served students in rising grades 1-8 (sometimes in addition to rising K or 9) across all schools in the district; two districts focused their programs on a smaller subset of grades and/or schools. Half of the programs were entirely voluntary. The other half used test scores and/or other criteria (e.g., course grades, attendance, teacher recommendations) to target summer school invitations, though these districts also generally allowed students who did not meet the criteria to participate. All districts hired summer school teachers from their incumbent workforce. One district hired a mix of their own teachers (~60%) and teachers from other districts, or teachers who were new to the district that summer (~40%). The maximum intended student-teacher ratio varied widely, from 1:10 to 1:25. One district operated a summer school program at each individual school, while the remaining districts operated hub sites that served students from multiple schools.

In addition to the information in Appendix Table A4, several characteristics were shared across all programs: providing professional development for teachers prior to the program, delivering grade-level instruction (based on a student's grade as of 2022-23) to address gaps in

learning from the prior year, and including some time for social emotional learning and/or enrichment activities (e.g., physical education, art, music, dance, gardening).

The intended duration and frequency (i.e., "dosage") varied across programs (see Appendix Table A4). About half the programs were open four days per week. The rest were open five days per week. The total days of programming ranged from 15 days to 20 days, although one district offered programs that ran for 23, 24, or 30 days, depending on the site. The programs ran for a total of 4 to 9 hours per day and generally included between 45 minutes and 2 hours of subject-specific instruction in math and reading. One program (District 2) provided 3 hours of literacy instruction per day to all participants in rising grades K-6 and allowed participants in rising grades 7-8 to choose to attend up to two total 3-hour sessions of math and/or literacy instruction. Excluding District 2's unique program, the total instructional hours offered in each subject ranged widely, from 11.25 hours up to 45 hours. In District 2, for students in rising grades 7-8 who chose to attend both 3-hour sessions in math or reading, the program provided a total of 108 hours of instruction in the relevant subject. Generally, our sample provided significantly less instruction than the recommended minimum of 75 hours (Schwartz et al., 2018) and fewer instructional days than the programs evaluated in the pre-pandemic literature (Augustine et al., 2016; Lynch et al., 2023).

3.2 Participation and Dosage in Summer 2023

Table 2 shows the percentage of students in eligible grades who participated in at least one day of summer school in each district. Across the eight districts, participation rates ranged from 3.0% to 23.0% (the overall average was 13.0%). Schools that served as "hub" sites had higher participation rates on average (22.8%) than non-hub schools (9.1%). District 7 had the highest overall participation rate at 23.0%, which was driven by particularly high participation

rates at hub sites (90.3%). Among the four districts that used specific criteria to target summer school invitations, the selectiveness of the targets varied substantially: three districts targeted between 11.0% and 18.5% of students in eligible grades, while the fourth targeted more than 50% of its students.

Across districts, the average number of days attended was 12.5, or 67% of the days offered. District attendance rates ranged from 59% to 75%. Program length and days of attendance were significantly correlated (r=0.920; p=0.001), and program length and attendance rates were not significantly correlated (r=0.459; p=0.252), suggesting that additional days of programming did not yield lower attendance rates. The total instruction students received averaged 17.7 hours in math (ranging from 7.3 to 24.5 hours) and 19.3 hours in reading (ranging from 7.3 to 34.9 hours).

Changes in Participation and Dosage from Summer 2022

As displayed in Appendix Table A5, the participation rate and total participation numbers in 2023 were lower on average than they were in 2022 in three of the five districts (where we have data for both years). The decreases in these three districts ranged from 3 to 5 percentage points. In the remaining two districts, participation increased by 2 and 3 percentage points respectively. The decreases in the three districts may be explained, at least in part, by two of the districts' reducing the number of grade-levels that could participate in summer school in 2023. Besides these participation drops, dosage also generally declined from 2022 to 2023. On average, daily attendance rates in 2023 were 4 percentage points lower than those of the same programs in 2022; 2023 participants also received 1.6 fewer hours of math instruction and 0.8 fewer hours of reading instruction compared to 2022 participants.

3.3 Academic Effects of Summers 2023 Programs

Panels A and B of Figure 1 (estimates reported in Appendix Tables A6 and A7) show the effect of attending at least one day of summer programming on fall 2023 MAP Growth scores respectively in math and reading. Preliminary results for RQ1 indicate that summer 2023 programs had a significant positive effect on math achievement on average (0.024 SD, $p < 0.01$), but this effect is small relative to the effects of pre-pandemic summer programs (+0.08-0.10 SD; Augustine et al., 2016; Lynch et al., 2022) and other education interventions (Kraft, 2020). Across districts, the estimated effects are positive and range in magnitude from 0.005 to 0.090 SD—four of the eight estimates are statistically significant.

We also calculate hourly effect estimates as the effect of any participation divided by the average instructional time dosage (in hours) for each program. This calculation allows us to compare program impacts per hour of instruction across sites. Among programs that had significant impacts, three of four have small estimated hourly impacts (ranging from 0.001 SD to 0.002 SD) that are statistically indistinguishable from each other. Although we cannot rule out similar impacts per hour of math instruction in these programs, all the average program effects are smaller than what we would expect based on prior estimates of summer school's impact on math achievement, which imply 0.005 SD gains per hour of instruction (McCombs et al., 2014), and the average program dosage in our sample (see [redacted] for methodology).

Results for reading were less positive. We did not detect a significant program impact on reading achievement on average (-0.008 SD, $p > 0.05$). In six of eight districts, we find effects that are generally small and statistically indistinguishable from zero, ranging from -0.023 SD to 0.044 SD. In districts 2 and 4, we find statistically significant negative impacts, -0.024 SD and -0.021 SD respectively. As we discuss in the conclusion, we hypothesize that the negative impacts

could reflect either subject-specific selection into summer programming (with lower-performing readers more likely to participate) or substantial reading instruction received by students in the comparison group/counterfactual. The hourly impacts of summer school on reading scores are all relatively small (-0.003 SD to 0.002 SD) and indistinguishable from zero for all districts.

Changes in Effects from Summer 2022

For the six districts in this study for which we have data on both summer 2023 and summer 2022 programming, we find no significant differences across years in the estimated impact of summer school on math or reading on average or for any individual district. Indeed, three of the six districts (Districts 3, 4, and 7) had remarkably similar impacts over the two summers, with differences ranging from just -0.022 SD to 0.015 SD across math and reading. These districts also maintained a significant, positive effect on math that ranged between 0.026 SD and 0.044 SD across both summers. The other three districts (Districts 1, 6, and 8) include one that had a statistically significant impact on math in 2023 but not in 2022, and two districts that did not have a significant impact on math either year. In reading, we see that District 4, which had a significant negative impact on reading in summer 2023 (-0.212 SD), had a similar estimated effect on reading in summer 2022 (-0.020 SD), though the 2022 estimate was not statistically significant. Meanwhile, District 8 had a significant and positive impact on reading achievement in summer 2022 (0.037 SD) that was similar in magnitude to its 2023 impact (0.034 SD), but the 2023 estimate was not statistically significant.

Subgroup Analyses

We also examine the extent to which summer program impacts varied by different student groups. Appendix Tables A8 and A9 show the differences in impacts for elementary and middle school students in math and reading, respectively. For both subjects, we find no clear pattern.

We also conducted analyses by student demographics, including race and ethnicity, gender, free and reduced price lunch status, special education status, English language learner (ELL) status, and baseline (i.e., spring 2023) MAP Growth percentile. (see Appendix Tables A10-A17). Small sample sizes limit our ability (i.e., statistical power) to detect significant differences between subgroups within districts, but the average differences in effects between subgroups are generally small and do not follow a clear pattern across districts. For example, in District 1, we find larger impacts of summer school on math for special education students and students who scored above the 50th percentile at baseline. By contrast, District 4's summer school had larger positive effects for students not receiving special education services; District 3's program had larger positive effects for students who scored below the 50th percentile at baseline.

Changes in Subgroup Effects (by Grade, Demographics) from Summer 2022

While the grade level results were mixed for 2023, our 2022 analyses revealed that the positive impacts observed in math were driven by elementary students. However, for subgroup analyses based on student demographics, the lack of a clear pattern in the findings is consistent with the 2022 results.

3.4 Differences in Summer 2023 Impacts by Summer 2022 Participation Status

Participation and Dosage

For RQ2, we examine the effects of attending summer school in 2023 in the four districts that had at least 100 students participate in summer school in 2022 and 2023 on two separate groups: first-timers (who did not attend summer school in 2022) and repeaters (who did attend in 2022). Table 3 describes the patterns of participation and dosage for both groups. First-time participants made up the majority (56% to 80%) of 2023 participants across all four districts,

while returning participants accounted for 20% to 44% of 2023 participants. In three of four districts, first-timers attended slightly less frequently than repeaters, resulting in their receiving fewer instructional hours (between 0.4 and 2.9 fewer hours in math and between 0.4 and 3.1 fewer hours in reading). In District 4, however, first-timers received an additional 2.9 and 2.5 hours of math and reading instruction, respectively.

Interaction Effects of Summer 2022 and 2023 Participation

We present the results of the interaction effect models in Panels A and B of Table 4 for math and reading respectively. The interaction effects, which capture the difference in the effect of attending summer school in 2023 for repeaters versus first timers, vary across the four districts, ranging from -0.031 SD to 0.053 SD in math and -0.046 SD to 0.052 SD in reading. The interaction effect is statistically significant in just one district—District 4—and is significant in math (0.040 SD) and reading (0.052 SD). This suggests the district's repeaters benefited more from summer school in 2023 than its first-timers did in both subjects. Across the three remaining districts, we cannot rule out similar impacts for repeaters and first-timers in math or reading. We note that over half of the estimated interaction effects for these districts are larger (greater than 0.03 SD in absolute value) than the average math and reading effects of attending summer school in 2023 estimated under RQ1. While the interaction effects vary in magnitude and direction across districts (and are estimated with limited precision due to the small sample size of repeaters), these results suggest that attending summer school for a second consecutive year can boost student achievement as much as—or even more than—attending for the first time.

4. Discussion

4.1 Summary

The pandemic-era summer programs in our sample consistently had a positive, albeit modest, impact on math achievement, but no detectable positive effect on reading. Program

effects on math (0.024 SD on average) were smaller than what pre-pandemic research has shown, both in overall effect and in gains per instructional hour. The modest magnitude of these effects notwithstanding, because summer school programs reach a substantial portion of students—13.0% of students in eligible grades on average—they hold promise for improving math achievement at scale.

The lack of positive impacts on reading are broadly consistent with results from this same sample of districts in summer 2022 as well as some pre-pandemic summer school research ([redacted]; Schwartz et al., 2018). While there is no definitive explanation for these null and negative reading effects, education researchers have observed that school-based factors tend to have a stronger impact on math than reading (e.g., Jackson et al., 2014; Riehl & Welch, 2022). One possible explanation is that math instruction happens primarily in classrooms, while reading skills are more likely to develop both at school and at home. Families and alternative summer learning programs may be more likely to support reading and language skill development in the summer and less likely to practice math. If that were the case, the difference in outcomes between summer school participants and non-participants would be smaller for reading than for math—it could even be negative, if the non-participants' reading instruction quality and/or dosage surpassed what participants received in summer school. To better understand the treatment-control contrast the impact estimate represents, future research should examine the instructional experiences of students who were eligible for (or invited to) summer school but did not attend to better understand the treatment-control contrast the impact estimates represent.

Despite the fact that achievement remains behind 2019 levels (NCES, 2025), the scale of summer programming in our sample narrowed in 2023 relative to 2022 (across the five districts in the sample both years). Average participation rates across rising grades 1 through 8 fell by 1.4

percentage points and attendance rates among participants decreased from 70% to 66%. Relative to the average 2022 participant, the average 2023 participant received 1.6 fewer hours of math instruction and 0.8 fewer hours of reading instruction. The total instructional hours students received (i.e., attended) in 2023 averaged 17.7 hours in math, ranging from 7.3 to 24.5 hours across districts, and averaged 19.3 hours in reading, ranging from 7.3 to 34.9 hours. Again, this instructional dose was well below the recommended 75 hours of instruction (Schwartz et al., 2018). While there are likely multiple reasons for these declines, two districts reported reducing the scale of their program in anticipation of the end of ESSER.

Our analysis of attending two consecutive years of summer school indicates that repeat participation does not generally yield diminishing returns. While dosage and attendance were slightly higher among repeaters than first-timers, the overall benefits of a second summer were not significantly larger than those of first-time participants in most districts. The findings broadly align with prior research (Augustine et al., 2016), suggesting each summer provides independent benefits. But we are not able to test whether the effects compound in a linear manner over time or partially (or wholly) fade out between summers. Further research is needed to assess whether cumulative exposure over multiple summers produces long-term gains.

4.2 Implications for Policy and Practice

Thanks to its strong pre-pandemic evidence base, high-dosage tutoring has been a popular academic recovery intervention following the pandemic. But a growing body of pandemic-era research suggests that summer school programs may offer a more scalable strategy for addressing learning loss (Carbonari et al., 2024a, 2024b; Kraft et al., 2024b). While efforts to expand tutoring programs during the pandemic have encountered implementation challenges and potentially diminishing returns, summer programs (despite more modest results) can reach

substantially more students with more reliable impacts. As districts weigh their recovery options, summer school's ability to deliver results at scale make it a promising approach for supporting district-wide recovery. In this study, we found that even programs delivering less than the recommended 75 hours of instruction per subject can yield positive impacts on math achievement.

Because our findings show positive effects for math but not reading, policymakers might be tempted to offer math-only summer programs. But we should interpret these results carefully. While the idea of focusing on math has some empirical support, our finding that summer school did not impact reading achievement suggests attending summer school did not improve reading *significantly more* than not attending summer school. For some of the reasons noted earlier, that could be an indication that students' scores are improving similarly in summer school and outside of summer school. We cannot know whether summer school participants would fall behind their non-participating peers in reading if summer school focused exclusively on math. Future research should experiment with different doses of math and reading instruction to see if replacing some or all reading instruction with additional math instruction negatively impacts reading achievement and/or positively impacts math achievement.

With the expiration of ESSER funds, districts must make strategic decisions about future academic recovery investments. Our findings suggest summer programs, despite their modest impacts, are a viable, reliable, and scalable option for accelerating student learning in math. Recovery still looms large—district-wide, these programs closed an average of 1-2 percent of the districts' remaining recovery as of spring 2022, similar to 2022 summer programs ([redacted]). Summer school will need to be paired with other strategies and, likely, systemic changes to education policy to address the persistent and increasingly unequal impacts of the pandemic on

achievement (Wycoff, 2025). To this end, future research should explore ways to enhance summer program efficacy at scale and investigate how participation across multiple summers influences students' long-term academic trajectories. Expanding participation and/or increasing program durations could increase the district-wide impact of summer school—a direction worth exploring in future policy and research.

References

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 Report No. RR-1557-WF). RAND Corporation.

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.

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*.

Barry, S. S., & Sass, T. R. (2022). *The impact of a 2021 summer school program on student achievement* (Georgia Policy Labs Report No. 40). Georgia State University. https://doi.org/10.57709/FAJ9-8597

Borman, G., Pyne, J., Pyne, J., & MGT Consulting (2024). *Summer Boost 2023 Evaluation Final Report*. Bloomberg Philanthropies. https://www.summerboost.org/wp-content/uploads/2024/06/Summer-Boost-2023-Evaluation-Final-Report.pdf

Carbonari, M. V., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Hashim, A. K., Kane, T. J., McEachin, A., Morton, E., Muroga, A., Patterson, T., & Staiger, D. O. (2024a). The impact and implementation of academic interventions during COVID-19: Evidence from the Road to Recovery Project. *AERA Open, 10.*

Carbonari, M. V., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Kane, T. J., McDonald, A., McEachin, A., Morton, E., Muroga, A., Salazar, A., & Staiger, D. O. (2024b). *Impacts of academic recovery interventions on student achievement in 2022–23* (CALDER Working Paper No. 303-0724). American Institutes for Research.

Chetty, R., 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). Using school choice lotteries to test measures of school effectiveness. *American Economic Review, 104*(5), 406–411.

Dewey, D., Fahle, E., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2024a). *Federal pandemic relief and academic recovery*. Center for Education Policy Research at Harvard University. https://cepr.harvard.edu/files/cepr/files/federal-pandemic-relief-and-academic-recovery-2024.pdf

Dewey, D., Kuhfeld, M., Fahle, E., Kane, T., & Reardon, S. F. (2024b, September). *Reconciling recent evidence on academic recovery in the US from 2022 to 2023*. NWEA.

https://www.nwea.org/research/publication/reconciling-recent-evidence-on-academic-recovery-in-the-us-from-2022-to-2023/

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 Report No. RR-A956-13). RAND Corporation. https://doi.org/10.7249/RRA956-13

DiMarco, B., & Jordan, P. W. (2022). *How local educators plan to spend billions in federal COVID aid*. FutureEd Explainer. 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 at Harvard University. https://cepr.harvard.edu/sites/hwpi.harvard.edu/files/cepr/files/explaining_covid_losses_5.23.pdf

Goldhaber, D. & Falken, G. (2025). *ESSER and student achievement: Assessing the impacts of the largest one-time federal investment in K12 schools* (CALDER Working Paper No. 301-0325-2). American Institutes for Research.

Guryan, J., Ludwig, J., Bhatt, M. P., Cook, P. J., Davis, J. M. V., Dodge, K., Farkas, G., Fryer, R. G., Mayer, S., Pollack, H., Steinberg, L., & Stoddard, G. (2023). Not too late: Improving academic outcomes among adolescents. *American Economic Review, 113*(3), 738–765. https://doi.org/10.1257/aer.20210434

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. *Annual Review of Economics*, *6*(1), 801-825.

Jacobson, L. (2024, October 10). White House plan yields 323K tutors, mentors to aid COVID learning recovery. *The 74*. https://www.the74million.org/article/white-house-plan-yields-323k-tutors-mentors-to-aid-covid-learning-recovery/

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.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241-253.

Kraft, M. A., Edwards, D. S., & Cannata, M. (2024a). *The scaling dynamics and causal effects of a district-operated tutoring program* (EdWorkingPaper No. 24-1030). Annenberg Institute at Brown University. https://doi.org/10.26300/zcw7-4547

Kraft, M. A., Schueler, B. E., & Falken, G. T. (2024b). *Scaling tutoring: Evidence from a meta-analysis of 265 randomized controlled trials* (EdWorkingPaper No. 24-955). Annenberg Institute at Brown University. https://edworkingpapers.com/ai24-1031

Lewis, K., & Kuhfeld, M. (2022). *Progress toward pandemic recovery: Continued signs of rebounding achievement at the start of the 2022–23 school year*. NWEA. https://www.nwea.org/research/publication/progress-toward-pandemic-recovery-2022

Lewis, K., & Kuhfeld, M. (2023). *Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery*. NWEA. https://www.nwea.org/research/publication/educations-long-covid-2023

Lewis, K., & Kuhfeld, M. (2024). *Recovery still elusive: 2023–24 student achievement highlights persistent achievement gaps and a long road ahead*. NWEA. https://www.nwea.org/research/publication/recovery-still-elusive-2024

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.

National Center for Education Statistics (NCES). (2025). *The Nation's Report Card: 2024 mathematics and reading assessments*. U.S. Department of Education.

NWEA (2019). MAP® Growth™ technical report. Portland, OR. https://www.nwea.org/uploads/2021/11/MAP-Growth-Technical-Report-2019_NWEA.pdf

Polikoff, M., Clay, I., & Silver, D. (2023). *Beyond test scores: Broader academic consequences of the COVID-19 pandemic on American students*. Center on Reinventing Public Education. https://crpe.org/beyond-test-scores-2023/

Reardon, S. F., Fahle, E. M., Ho, A. D., Shear, B. R., Saliba, J., Min, J., Shim, J., & Kalogrides, D. (2025). Stanford Education Data Archive (Version SEDA 2024). Retrieved from: https://purl.stanford.edu/pt329xg7054.

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.

Roche, D. (2023, June 20). 75+ ESSER III spending categories. *Burbio school tracker.* https://about.burbio.com/weekly-updates/burbio-school-tracker-6/20-75-esser-iii-spending-categories

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 Report No. RR366-1). RAND Corporation. https://www.rand.org/pubs/research_reports/RR366-1.html

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

Wyckoff, J. H. (2025). *Puzzling over declining academic achievement* (EdWorkingPaper: 25-1197). Annenberg Institute at Brown University. https://doi.org/10.26300/kjaa-ec47.

Tables and Figures

Table 1. Sample Demographics

	[redacted] district students	U.S. public school students
Average district enrollment	50,749	2,720
Average school enrollment	620	397
FRPL eligible (%)	75.5%	54.0%
Race (%)		
Asian	3.3%	5.6%
Hispanic	45.4%	29.4%
Black	23.3%	14.9%
White	22.2%	43.7%
School locale (%)		
City	87.2%	31.0%
Suburb	7.9%	38.6%
Town	0.0%	10.2%
Rural	4.9%	20.1%

Note. FRPL = free-or-reduced-priced lunch. Data are from the Common Core of Data collected by the National Center for Education Statistics during the 2023-24 school year. Percentages are weighted by student enrollment. For states that do not report free-or-reduced price lunch data, direct certification counts were used in their place.

Table 2. Summer School Participation and Dosage in 2023

District	Sample rising grade levels	Sample size	% Targeted	% Treated			Intended program length (in days)	Average days attended	% Total days attended	Total hours of instruction per participant	
				Overall	Hub sites	Not hub				Math	Reading
1	7-8	7,759	18.5%	3.0%	1.7%	1.2%	20	12.2	61%	24.5	24.5
2	1-6	14,473	N/A	24.0%	N/A	N/A	18	13.8	77%	10.4	41.5
	7-8	5,466	14.5%	15.5%	N/A	N/A	18	12.3	68%	23.2	17.3
3	1-4	16,590	52.7%	12.0%	13.4%	10.6%	19	12.7	67%	25.4	25.4
	6-8	12,826	66.4%	7.8%	9.3%	6.3%	19	11.6	61%	11.6	17.4
4	1-8	50,915	N/A	11.5%	17.9%	7.9%	15	10.2	68%	20.5	19.8
5	1-8	11,256	N/A	20.5%	21.5%	19.1%	23, 24, 30[a]	18.5	75.0%	22.5	27.7
6	3-5	3,742	N/A	6.4%	7.5%	6.3%	15	9.5	63.5%	7.1	7.1
	6-8	3,368	N/A	4.7%	3.1%	5.3%	15	8.2	54.4%	-	-
7	1-8	22,195	N/A	23%	90.3%	13.0%	17	11.3	67%	17.0	17.0
8	1-8	28,271	11.0%	8.4%	10.8%	7.9%	20	14.2	71%	12.4	16.6

[a]Most of District 5's programs were 23 days long, but program length varied by site. Programs were 23, 24, or 30 days.
Note. "–" = data not available. With the exception of Districts 1 and 6, this table reports participation and dosage for each district's main summer program based on the total number of students enrolled in the district in grades 1-8 who were eligible for summer programming, regardless of whether the students had a valid NWEA MAP Growth score. For District 1, students in rising grade 4 were also eligible for programming but were excluded from the sample because the district did not provide data for this grade. For District 6, students in rising grades 1 and 2 were also eligible for programming but were excluded from the sample because the district did not provide data for these grades. District 6's hours of instruction per participant are unknown because the district did not provide information about subject-specific instructional time.

Table 3. Participation and Dosage for Summer School 2023 First-Timers and Repeaters

			Any Summer 2023 (First-timers and Repeaters)			First-timers (SS23 only)			Repeaters (SS22 and SS23)		
				Average Hours of Instruction			Average Hours of Instruction			Average Hours of Instruction	
District	Grades in Sample	Full Sample Size	% of Sample	Math	Reading	% of Sample	Math	Reading	% of Sample	Math	Reading
District 3	2-4, 6-8 (math) 4, 6-8 (reading)	25,430	10.4%	21.2	24.4	6.7%	20.2	23.3	3.7%	23.1	26.4
District 4	2-8	44,697	11.7%	19.8	19.2	9%	20.4	19.8	2.6%	17.5	17.3
District 7	1-5	11,407	32.0%	17.6	17.6	18.0%	16.4	16.4	14.1%	19.1	19.1
District 8	4-8	13,415	6.0%	12.9	12.9	4.8%	12.8	12.8	1.2%	13.2	13.2

Note. Average hours of instruction reflects the average dosage of all summer programming a student received in Summer 2023. For Districts 3 and 4, this includes tutoring as well as traditional summer school.

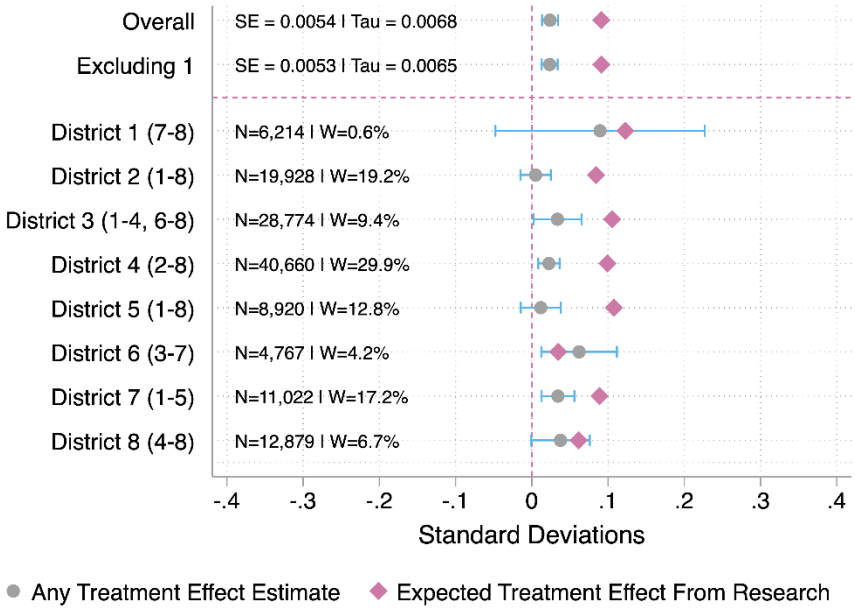
Table 4. Interaction Effects of Participating in Summer School in 2023 and 2022

	Panel A: Math					Panel B: Reading				
District	Rising Grades	Any SS 22 (SE)	Any SS 23 (SE)	Any SS 22 x Any SS 23 (SE)	N	Rising Grades	Any SS 22 (SE)	Any SS 23 (SE)	Any SS 22 x Any SS 23 (SE)	N
Overall		-0.0198*	0.0305**	0.0159	89,349		-0.0051	0.0012	0.0073	71,819
		(0.0062)	(0.0086)	(0.0179)			(0.0144)	(0.0145)	(0.0198)	
District 3	2-4, 6-8	-0.0311*	0.0516**	-0.0309	24,788	4, 6-8	-0.0311*	0.0110	-0.0117	16,537
		(0.0127)	(0.0172)	(0.0255)			(0.0142)	(0.0210)	(0.0264)	
District 4	2-8	-0.0243**	0.0167*	0.0398*	40,660	2-8	-0.0200	-0.0313**	0.0517*	31,920
		(0.0090)	(0.0079)	(0.0160)			(0.0107)	(0.0113)	(0.0218)	
District 7	1-5	-0.0095	0.0347*	0.0051	11,022	1-5	0.0024	0.0099	0.0017	10,325
		(0.0160)	(0.0139)	(0.0234)			(0.0180)	(0.0157)	(0.0269)	
District 8	4-8	0.0064	0.0376	0.0534	12,879	4-8	0.0371*	0.0336	-0.0456	13,037
		(0.0174)	(0.0196)	(0.0368)			(0.0184)	(0.0250)	(0.0448)	

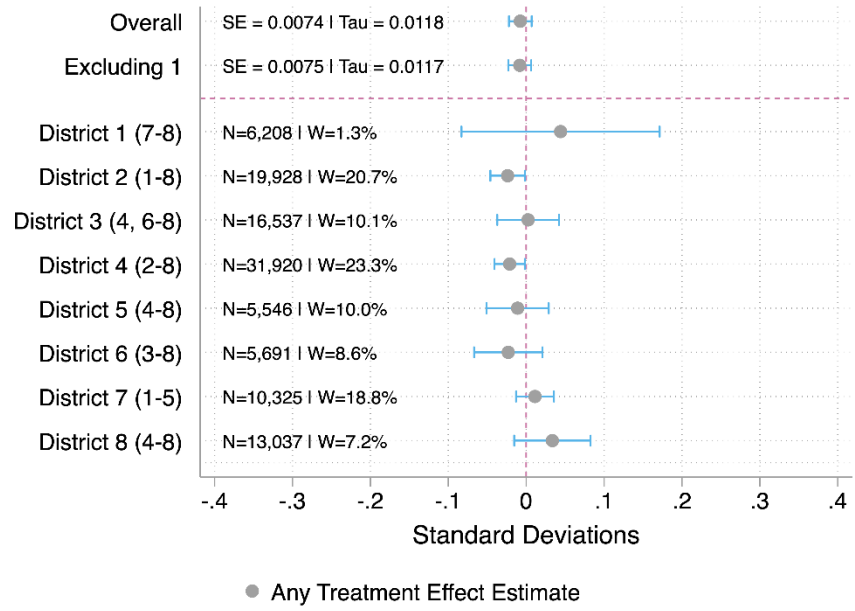
Note. The "Overall" estimate refers to the meta-analytic estimates of the four coefficients for each subject. In each panel, the analytic sample includes all students entering Grades 1 through 8 in fall 2023 who had a valid spring 2023 and fall 2023 MAP Growth test score in the relevant subject. For District 8, where spring 2023 MAP Growth testing was very limited, the analytic sample includes students who had a non-missing winter 2023 MAP Growth score and spring 2023 state standardized test score in addition to the fall 2023 MAP Growth score in the relevant subject.
*p<0.05, **p<0.01

Figure 1. Effects of Attending Summer School in 2023 on Fall MAP Growth Test

Panel A: Math



Panel B: Reading



Note. The blue bars represent the 95% confidence interval. For math, the expected treatment effect from research is based on a 0.11 SD increase in math achievement associated with attending an average of 21.9 hours of math instruction in a summer program (McCombs et al., 2014).

Appendix

Appendix Table A1. Estimated Achievement Loss and Recovery from Spring 2019 to 2024, Grades 3-8

	District	Spring 2019 (SDs)	Spring 2022 (SDs)	Spring 2023 (SDs)	Spring 2024 (SDs)	Change from 2019 to 2022 (SDs)	Change from 2022 to 2024 (SDs)	Change from 2019 to 2024 (SDs)
Panel A: Math	Alexandria	–	–	–	–	–	–	–
	Anonymous[a]	–	–	–	–	–	–	–
	Dallas	-0.08	-0.21	-0.14	-0.24	-0.13	-0.03	-0.17
	Guilford	-0.10	-0.22	-0.23	-0.19	-0.12	0.03	-0.09
	Portland	0.01	–	–	–	–	–	–
	Santa Ana	-0.51	-0.63	-0.62	-0.61	-0.13	0.02	-0.11
	Syracuse	–	–	–	–	–	–	–
	Tulsa	-0.65	-1.08	-1.13	-1.12	-0.43	-0.04	-0.47
	Study District Avg	-0.27	-0.54	-0.53	-0.54	-0.20	-0.01	-0.21
	National District Avg	0.00	-0.14	-0.12	-0.11	-0.14	0.03	-0.11
Panel B: Reading	Alexandria	-0.14	-0.42	-0.46	-0.47	-0.28	-0.04	-0.33
	Anonymous[a]	–	–	–	–	–	–	–
	Dallas	-0.25	-0.33	-0.34	-0.37	-0.08	-0.04	-0.12
	Guilford	-0.03	-0.19	-0.25	-0.29	-0.16	-0.10	-0.26
	Portland	0.16	–	–	–	–	–	–
	Santa Ana	-0.53	-0.53	-0.57	-0.62	0.00	-0.09	-0.09
	Syracuse	–	–	–	–	–	–	–
	Tulsa	-0.58	-0.95	-0.97	-0.93	-0.37	0.02	-0.35
	Study District Avg	-0.25	-0.50	-0.53	-0.55	-0.15	-0.05	-0.21
	National District Avg	0.02	-0.06	-0.09	-0.11	-0.08	-0.05	-0.13

[a]Enrollment/racial composition changed significantly in this district from 2019 to 2022 and results should be interpreted with caution.
Note. All estimates are from the Stanford Education Data Archive (Version SEDA 2024; Reardon et al., 2025) and are scaled such that a 0 in this metric is equal to the average of the national NAEP average (in grade 5.5) in spring 2019, and 1 unit in this metric is equal to 1 student level standard deviation (SD). Estimates in this scale are comparable across the whole country, and over time, but they are not comparable across subjects. "–" indicates the relevant estimate was not available in SEDA.

Appendix Table A2. Sample Characteristics by Treatment Status in Summer 2023

	District 1		District 2		District 3		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	12.2%	6.0%	15.1%	16.3%	22.4%	10.6%	8.1%	6.2%	21.1%	20.6%	6.8%	8.0%	13.0%	9.7%	24.86%	19.75%
ELL	8.2%	7.8%	55.3%	43.5%	15.4%	7.9%	59.4%	53.1%	16.6%	19.5%	56.5%	38.8%	30.7%	38.1%	25.14%	13.25%
FRPL	–	–	75.3%	75.6%	–	–	–	–	84.3%	83.2%	–	–	43.5%	45.1%	–	–
Male	53.2%	51.7%	50.9%	50.7%	52.2%	50.7%	50.4%	50.5%	51.3%	50.8%	55.9%	51.5%	50.9%	51.6%	41.76%	41.01%
Race:																
Black	66.7%	27.1%	0.16%	0.18%	53.5%	40.0%	16.5%	14.2%	58.4%	42.4%	37.6%	23.4%	22.5%	19.7%	13.36%	6.13%
Hispanic	8.7%	12.6%	96.6%	95.9%	25.6%	18.0%	77.7%	74.6%	12.0%	16.8%	27.5%	37.4%	33.2%	40.7%	20.81%	13.8%
Asian	2.6%	1.8%	1.5%	1.9%	5.0%	7.3%	1.2%	1.1%	5.5%	7.7%	12.1%	7.5%	1.9%	1.7%	5.68%	4.38%
Other	2.2%	3.1%	0.79%	0.99%	5.5%	5.7%	1.6%	2.2%	9.8%	11.1%	2.7%	2.8%	17.6%	17.0%	12.43%	11.09%
White	19.9%	55.4%	0.69%	0.84%	10.4%	29.0%	3.0%	7.9%	14.3%	22.1%	20.1%	28.9%	24.8%	20.9%	26.30%	45.52%
Priority for summer	65.8%	21.2%	68.9%	65.9%	90.5%	55.1%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Spring 2023 MAP test[a]																
Math (RIT points)	213.9	227.8	185.0	196.2	178.1	197.0	194.8	203.46	185.1	183.2	193.9	203.8	175.7	175.5	184.04	204.02
Math (normed)	-0.59	0.17	-0.65	-0.51	-1.29	-0.17	-0.63	-0.14	-1.05	-0.94	-0.90	-0.42	-0.78	-0.78	-0.15	0.13
Reading (RIT points)	207.9	219.6	184.2	195.4	179.4	200.6	187.6	198.5	188.1	187.6	187.0	199.8	172.0	172.5	179.40	201.63
Reading (normed)	-0.54	0.17	-0.69	-0.46	-1.36	-0.17	-0.91	-0.26	-1.09	-0.99	-0.96	-0.32	-0.72	-0.70	-0.14	0.18
Fall 2023 MAP test																
Math (RIT points)	213.5	227.3	184.0	195.1	180.9	198.5	193.1	199.39	185.4	183.18	195.6	203.2	173.5	172.8	191.72	211.92
Math (normed)	-0.49	0.26	-0.66	-0.49	-0.98	0.02	-0.61	-0.18	-1.03	-1.01	-0.57	-0.16	-0.92	-0.95	-0.29	0.22
Reading (RIT points)	207.3	219.1	181.6	192.9	187.5	207.6	188.6	198.1	187.0	186.8	190.2	202.6	171.7	171.7	188.26	210.45
Reading (normed)	-0.51	0.18	-0.70	-0.45	-0.63	-0.01	-0.78	-0.22	-0.96	-0.86	-0.71	-0.12	-0.77	-0.77	-0.19	0.37
N	231	6,038	4,323	15,605	2,991	26,425	5,224	45,041	2,089	6,975	338	5,511	3,655	7,752	2,148	26,123

[a]For District 8, we report Winter 2023 scores instead of Spring 2023 scores because of low testing rates in the spring.

Note. "Special Education" includes all students who have an Individualized Education Plan (IEP) or a 504 plan; for District 1, only 504 plan status was available. ELL = English language learner. FRPL = free or reduced-priced lunch. Student-level FRPL data were not available in Districts 1, 3, 4, 6, and 8. "Priority for summer" refers to students who were prioritized or targeted to receive summer programming based on criteria that varied by district. Districts 4, 5, 6, 7, and 8 did not prioritize subsets of students to participate.

Appendix Table A3. Sample Characteristics by Treatment Status across Summers 2022 and 2023

	District 3				District 4				District 7				District 8			
	No summer	SS22 only	First-timers (SS23 only)	Repeaters (SS22 & SS23)	No summer	SS22 only	First-timers (SS23 only)	Repeaters (SS22 & SS23)	No summer	SS22 only	First-timers (SS23 only)	Repeaters (SS22 & SS23)	No summer	SS22 only	First-timers (SS23 only)	Repeaters (SS22 & SS23)
Special Education	9.5%	23.1%	19.8%	28.2%	10.1%	17.2%	13.9%	22.2%	9.4%	11.7%	12.2%	14.0%	16.6%	38.4%	24.8%	42.0%
ELL	8.5%	16.6%	15.0%	21.6%	52.6%	59.5%	58.7%	61.9%	38.5%	35.5%	30.7%	30.6%	12.0%	19.9%	31.6%	24.4%
FRPL	-	-	-	-	80.8%	88.4%	87.4%	90.6%	45.7%	40.9%	48.8%	36.7%	-	-	-	-
Male	50.3%	53.8%	52.4%	50.5%	50.5%	50.2%	50.5%	50.1%	51.5%	52.7%	51.9%	49.7%	50.7%	52.2%	55.9%	51.1%
Race:																
Black	38.8%	53.0%	55.2%	50.3%	19.5%	16.7%	24.7%	16.3%	19.5%	21.1%	23.9%	20.6%	5.7%	16.8%	15.8%	26.1%
Hispanic	17.6%	23.6%	23.9%	28.7%	81.4%	79.6%	85.2%	80.0%	41.1%	38.3%	33.4%	32.9%	15.1%	29.3%	28.7%	29.5%
Asian	7.6%	4.9%	5.0%	4.9%	1.7%	0.5%	1.9%	1.0%	1.8%	1.5%	2.1%	1.7%	5.1%	4.6%	10.3%	2.3%
Other	5.7%	5.3%	5.6%	5.4%	2.3%	1.2%	1.2%	0.9%	17.2%	15.9%	18.2%	16.9%	13.0%	16.3%	15.5%	15.9%
White	30.3%	13.3%	10.1%	10.7%	11.9%	2.2%	5.3%	2.0%	20.5%	23.2%	22.4%	27.8%	61.1%	33.0%	29.6%	26.1%
Priority for summer	53.3%	85.8%	90.4%	89.1%	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Spring 2023 MAP test[a]																
Math (RIT points)	204.90	188.78	184.42	179.98	203.12	206.68	192.96	201.05	174.67	180.35	172.86	179.35	216.16	201.39	196.32	198.67
Math (normed)	-0.11	-1.00	-1.26	-1.37	-0.11	-0.58	-0.58	-0.83	-0.79	-0.69	-0.89	-0.64	0.15	-0.32	-0.35	-0.37
Reading (RIT points)	204.78	188.96	183.56	177.53	198.58	197.73	186.52	191.05	171.81	176.78	169.40	175.45	215.92	201.36	195.18	194.35
Reading (normed)	-0.09	-0.99	-1.28	-1.51	-0.21	-0.80	-0.82	-1.17	-0.70	-0.70	-0.80	-0.63	0.22	-0.13	-0.24	-0.37
Fall 2023 MAP test																
Math (RIT points)	206.08	190.09	187.45	181.92	198.94	203.41	191.14	199.93	172.09	177.58	170.78	176.88	218.33	201.71	198.87	199.09
Math (normed)	0.10	-0.78	-0.91	-1.09	-0.16	-0.55	-0.55	-0.73	-0.97	-0.85	-1.03	-0.79	0.29	-0.81	-0.69	-0.91
Reading (RIT points)	209.68	193.78	190.30	183.95	198.15	198.17	187.18	193.04	171.00	176.45	168.91	175.30	211.14	199.98	190.28	190.16
Reading (normed)	0.05	-0.55	-0.66	-0.76	-0.19	-0.66	-0.63	-0.87	-0.78	-0.71	-0.86	-0.66	0.44	-0.52	-0.59	-0.86
N	20,852	1,930	1,705	943	36,643	2,830	4,044	1,180	6,682	1,070	2,049	1,606	12,693	719	658	176

[a]For District 8, we report Winter 2023 scores instead of Spring 2023 scores because of low testing rates in the spring.

Note. "Special Education" includes all students who have an Individualized Education Plan (IEP) or a 504 plan. ELL = English language learner. FRPL = free or reduced-priced lunch. Student-level FRPL data were not available in Districts 3 and 8. "Priority for summer" refers to students who were prioritized or targeted to receive summer programming based on criteria that varied by district. Districts 4, 7, and 8 did not prioritize subsets of students to participate.

Appendix Table A4. Designs of Summer School Programs Serving Students in Rising Grades K-9

Panel A: Districts 1-4

	District 1	District 2	District 3	District 4
Grade levels	Rising grades 4, 7-9	Rising grades K-8	Rising grades 1-4, 6-9	Rising grades K-8
Eligible schools	All schools	All schools	Grades 1-2, 6-9: All schools Grades 3-4: Title I schools	Reg. calendar schools
Participation: Opt-in or by invitation?	Grade 4: Invitation only Grades 7-9: Invitation, but opt-in allowed	Grades K-6: Opt-in Grades 7-8: Invitation, but opt-in allowed	Invitation, but opt-in allowed	Opt-in
Invitation criteria	Grade 4: Below grade-level on state test Grades 7-9: 20th-35th percentile on MAP math or 20th-41.5th percentile on MAP reading, other test scores, course grades	Grades K-6: All students Grades 7-8: Below 50th percentile on MAP math (for math sessions) or reading (for reading sessions)	Below grade-level on state tests or DIBELS, or below 40th percentile on MAP in math or reading; absenteeism; teacher recommendations	All students
Academic providers	District teachers and teaching assistants	District teachers	District teachers, teaching assistants, and tutors	District teachers
Student-teacher ratio	1:15 max	1:20 max	1:10 max	1:20 max
Location	Hub sites	All school sites	Hub sites	Hub sites
Intended frequency	Grade 4: 5 days/wk Grades 7-9: 4 days/wk	5 days/wk	4 days/wk	4 days/wk
Intended dosage	Grade 4: 15 days Grades 7-9: 20 days	18 days	19 days	15 days
Math instruction per day	2 hrs	Grades K-6: 3 hrs Grades 7-8: 0, 3, or 6 hrs	Grades 1-4, 6: 2 hrs Grades 7-9: 1 hr	2 hrs
Reading instruction per day	2 hrs	Grades K-6: 45 min Grades 7-8: 0, 3, or 6 hrs	Grades 1-4, 6: 2 hrs Grades 7-9: 1 hr	2 hrs
Operating hours	Grade 4: 6 hrs Grades 7-9: 4.5 hrs	Grades K-6: 4.5 or 9 hrs Grades 7-8: 3, 6.5, or 9 hrs	Grades 1-4, 6: 6.75 hrs Grades 7-9: 5.25 hrs	Grades K-5: 8 hrs Grades 6-8: 8 hrs
Other programming			Tutoring, enrichment camps, HS programming	Enrichment camps, HS programming

Panel B: Districts 5-8

	District 5	District 6	District 7	District 8
Grade levels	Rising grades 1-9	Rising grades 1-9	Rising grades K-9	Rising grades 1-8
Eligible schools	All schools	All schools	All schools	All schools
Participation: Opt-in or by invitation?	Opt-in	Opt-in	Opt-in	Invitation, but opt-in allowed
Invitation criteria	All students	All students	All students	5th-40th or 10th-40th percentile on MAP reading (depending on site), other test scores, teacher recommendations
Academic providers	District teachers	District teachers	District teachers	~60% district teachers, ~40% teachers from other districts or new teachers
Student-teacher ratio	1:25 max	1:18 max	1:25 max	1:20 max
Location	Hub sites	Hub sites	Hub sites	Hub sites
Intended frequency	4 or 5 days/wk, depending on site	4 days/wk	5 days/wk	5 days/wk
Intended dosage	23, 24, or 30 days, depending on site	15 days	18 days	20 days
Math instruction per day	Grades 1-6: 1 hr Grades 7-9: 1.5 hrs	Grades 1-5: 45 min Grades 6-9: unknown	1.5 hrs	Grades 1-2: 45 min Grades 3-5: 1 hr Grades 6-8: 50 min
Reading instruction per day	Grades 1-6: 1.5 hrs Grades 7-9: 1.5 hrs	Grades 1-5: 45 min Grades 6-9: unknown	1.5 hrs	Grades 1-2: 1.5 hrs Grades 3-5: 1 hr Grades 6-8: 50 min
Operating hours	6 hrs	Grades 1-5: 4 hrs Grades 6-9: 4.25 hrs	7.5 hrs	4 hrs
Other programming	Enrichment camps, HS programming	Enrichment camps, HS programming	Tutoring	Tutoring, enrichment camps

Notes. HS=high school (rising grades 10-12). For the purposes of this table, we consider all literacy, English Language Arts, and reading instruction to be "reading instruction." At District 2, students in rising grades 7-8 could choose whether to participate in 1 or 2 total 3-hour sessions that were focused on math or ELA. For District 3, instructional times were not available for summer 2024. With support from the district, we assumed the times in 2024 were the same as 2023.

Appendix Table A5. Changes in Summer School Participation and Dosage from Summer 2022 to Summer 2023

District	Summer 2022						Summer 2023					
	# Students targeted (% of grades 1-8)	# Students participated (% of grades 1-8)	Average days attended	% Total days attended	Total hours of instruction per participant		# Students targeted (% of grades 1-8)	# Students participated (% of grades 1-8)	Average days attended	% Total days attended	Total hours of instruction per participant	
					Math	Reading					Math	Reading
1	–	–	13.6	80%	27.2	27.2	–	–	12.2	61%	24.5	24.5
2	–	–	–	–	–	–	13,260 (53%)	4,327 (17%)	13.5	75%	12.9	36.8
3	18,488 (55%)	4,037 (12%)	10.9	65%	20.2	21.0	20,610 (49%)	3,447 (8%)	12.3	65%	20.8	22.7
4	N/A	5,881 (5%)	10.8	72%	16.2	16.2	N/A	5,874 (7%)	10.2	68%	19.8	19.2
5	–	–	–	–	–	–	N/A	2,302 (21%)	18.5	75%	22.5	27.7
6[a]	N/A	849 (11%)	9.5	63%	14.2	14.2	N/A	398 (6%)	9.0	64%	6.7	6.7
7	N/A	4,417 (20%)	12.5	66%	18.7	18.7	N/A	5,109 (23%)	11.3	67%	17.0	17.0
8	10,021 (44%)	2,523 (11%)	14.2	71%	14.2	14.2	3,095 (11%)	2,374 (8%)	14.2	71%	12.4	16.6

[a]Enrollment data for District 6 were limited to grades 3-8 in both years. The total instructional hours per participant in Summer 2023 for District 6 are calculated by assuming students in rising grades 6-8 received the same amount of instructional time in math and reading per day as the students in rising grades 3-5 because daily instructional time for students in rising grades 6-8 was unknown.

Note. The denominator for the percentage of students targeted and the percentage of students participated includes all students enrolled in the district in rising grades 1-8. The numerators include all students enrolled in the district who were targeted or participated in summer programming in rising grades 1-8, regardless of whether the students had a valid NWEA MAP Growth score. Data on enrollment across all of grades 1-8 were not available for District 1 in either year.

Appendix Table A6. Effects of Attending Summer School on Fall MAP Growth Test, Math

	(A) Summer 2023							(B) Summer 2022		Difference (A) - (B)
District (Rising Grades)	Any SS 23 (SE)	Hourly SS 23 (SE)	Avg Dosage in Days	Avg Dosage in Hours	Expected Effect from Research	% Treated	N	Any SS 22 (SE)	Hourly SS 22 (SE)	Any SS 23 - Any SS 22 (SE)
Overall	0.0238** (0.0054)	0.0012** (0.0002)	12.60	18.18	0.0913	15.41%	133,164	0.0259** (0.0068)	0.0014** (0.0004)	-0.0021 (0.0087)
Overall (omitting District 1)	0.0234** (0.0053)	0.0012** (0.0003)	12.60	18.11	0.0909	15.98%	126,950	0.0314** (0.0049)	0.0018** (0.0003)	-0.0080 (0.0072)
District 1 (7-8)	0.0895 (0.0700)	0.0037 (0.0029)	12.23	24.46	0.122	3.72%	6,214	-0.0112 (0.0150)	-0.0004 (0.0005)	0.1007 (0.0716)
District 2 (1-8)	0.0049 (0.0102)	0.0003 (0.0006)	13.06	16.79	0.084	21.69%	19,928	N/A	N/A	N/A
District 3[a] (1-4, 6-8)	0.0336** (0.0161)	0.0016** (0.0008)	12.39	20.95	0.105	10.12%	28,774	0.0339** (0.0115)	0.0018** (0.0006)	-0.0003 (0.0198)
District 4 (2-8)	0.0224** (0.0071)	0.0011** (0.0004)	9.90	19.8	0.099	12.17%	40,660	0.0267** (0.0081)	0.0017** (0.0005)	-0.0043 (0.0108)
District 5 (1-8)	0.0118 (0.0134)	0.0006 (0.0006)	18.64	21.41	0.107	23.12%	8,920	N/A	N/A	N/A
District 6 (3-7)	0.0620* (0.0253)	0.0090*[b] (0.0037)	9.17	6.88[b]	0.035	6.38%	4,767	0.0181 (0.0247)	0.0012 (0.0017)	0.0439 (0.0354)
District 7 (1-5)	0.0342** (0.0110)	0.0019** (0.0006)	11.79	17.69	0.088	32.25%	11,022	0.0435** (0.0105)	0.0023** (0.0006)	-0.0093 (0.0152)
District 8 (4-8)	0.0376 (0.0196)	0.0032 (0.0015)	14.18	12.26	0.061	17.00%	12,879	0.0064 (0.0174)	0.0004 (0.0012)	0.0312 (0.0261)

[a]In 2022, the grade range for this district was 1-8. In 2023, rising fifth graders were ineligible for summer programming, so the analysis sample is restricted to rising grades 1-4 and 6-8.

[b]The hourly effect, average dosage in hours, and expected effect for District 6 are calculated by assuming students in rising grades 6-7 received the same amount of instructional time in math and reading per day as the students in rising grades 3-5 because daily instructional time for students in rising grades 6-8 was unknown.

Note. The "Overall" estimate refers to the meta-analytic estimates of the eight coefficients. The second overall estimate excludes District 1 because the district's analysis sample included only middle school grades with low participation rates, which tended to have lower effects overall. The expected treatment effect from research is based on a 0.11 SD increase in math achievement associated with attending an average of 21.9 hours of math instruction in a summer program (McCombs et al., 2014). *p<0.05, **p<0.01

Appendix Table A7. Effects of Attending Summer School on Fall MAP Growth Test, Reading

	(A) Summer 2023							(B) Summer 2022		Difference (A) - (B)
District (Rising Grades)	Any SS 23 (SE)	Hourly SS 23 (SE)	Avg Dosage in Days	Avg Dosage in Hours	Expected Effect from Research	% Treated	N	Any SS 22 (SE)	Hourly SS 22 (SE)	Any SS 23 - Any SS 22 (SE)
Overall	-0.0076 (0.0074)	-0.0004 (0.0003)	12.45	21.85	N/A	15.27%	109,192	0.0047 (0.0076)	0.0002 (0.0004)	-0.0123 (0.0106)
Overall (omitting District 1)	-0.0082 (0.0075)	-0.0005 (0.0003)	12.45	21.81	N/A	15.97%	102,984	0.0066 (0.0078)	0.0003 (0.0004)	-0.0148 (0.0108)
District 1 (7-8)	0.0442 (0.0649)	0.0018 (0.0027)	12.23	24.46	N/A	3.72%	6,208	-0.0351 (0.0327)	-0.0013 (0.0012)	0.0793 (0.0727)
District 2 (1-8)	-0.0238* (0.0113)	-0.0008 (0.0004)	13.06	29.39	N/A	21.69%	19,928	N/A	N/A	N/A
District 3[a] (4, 6-8)	0.0025 (0.0202)	0.0001 (0.0011)	12.17	20.87	N/A	9.49%	16,537	0.0225 (0.0135)	0.0012 (0.0007)	-0.0220 (0.0243)
District 4 (2-8)	-0.0212* (0.0099)	-0.0011 (0.0005)	9.61	19.22	N/A	10.97%	31,920	-0.0197 (0.0132)	-0.0013 (0.0009)	-0.0015 (0.0165)
District 5 (4-8)	-0.0110 (0.0204)	-0.0004 (0.0009)	18.43	27.65	N/A	24.09%	5,546	N/A	N/A	N/A
District 6 (3-8)	-0.0230 (0.0224)	-0.0034[b] (0.0033)	9.01	6.76[b]	N/A	5.71%	5,691	0.0200 (0.0222)	0.0014 (0.0015)	-0.0430 (0.0315)
District 7 (1-5)	0.0113 (0.0124)	0.0006 (0.0007)	11.61	17.42	N/A	32.04%	10,325	-0.0032 (0.0118)	-0.0002 (0.0006)	0.0145 (0.0171)
District 8 (4-8)	0.0336 (0.0250)	0.0015 (0.0014)	14.19	16.747	N/A	16.00%	13,037	0.0371* (0.0184)	0.0026 (0.0013)	-0.0035 (0.031)

[a]In 2022, the grade range for this district was 1-8. In 2023, rising fifth graders were ineligible for summer programming, so the analysis sample is restricted to rising grades 1-4 and 6-8.

[b]The hourly effect and average dosage in hours for District 6 are calculated by assuming students in rising grades 6-8 received the same amount of instructional time in math and reading per day as the students in rising grades 3-5 because daily instructional time for students in rising grades 6-8 was unknown.

Note. The "Overall" estimate refers to the meta-analytic estimates of the eight coefficients. The second overall estimate excludes District 1 because the district's analysis sample included only middle school grades with low participation rates, which tended to have lower effects overall. The hourly effect and average dosage for District 6 are calculated by assuming students in rising grades 6-8 received the same amount of instructional time in math and reading per day as the students in rising grades 3-5 because daily instructional time for students in rising grades 6-8 was unknown. *p<0.05, **p<0.01

Appendix Table A8. Effects of 2023 Summer School on Fall MAP Growth Test by Grade Bands, Math

	(A) Rising Grades 1-5					(B) Rising Grades 6-8					Difference (A) - (B)
District	Grades in Sample	Any SS 23 (SE)	Average Dosage (Hours)	% Treated	N	Grades in Sample	Any SS 23 (SE)	Average Dosage (Hours)	% Treated	N	Any SS 23 (SE)
1	N/A	N/A	N/A	N/A	N/A	7-8	0.0895 (0.0700)	24.46	3.72%	6,214	N/A
2	1-5	0.0049 (0.0102)	10.39	24.20%	14,371	7-8	0.0130 (0.0354)	23.22	15.82%	5,366	-0.0081 (0.0368)
3	1-4	0.0546** (0.0246)	25.41	11.94%	16,445	6-8	0.0018 (0.0161)	11.72	7.69%	12,329	0.0532 (0.0294)
4	2-5	0.0253** (0.0087)	20.01	12.17%	26,334	6-8	0.0168 (0.0122)	19.29	12.17%	14,326	0.0085 (0.015)
5	1-5	-0.0012 (0.0195)	19.24	20.29%	5,578	6-8	0.0292 (0.0170)	24.06	27.83%	3,342	-0.0304 (0.0259)
6	3-5	0.0656 (0.0325)	7.14	6.73%	3,106	N/A	N/A	N/A	N/A	N/A	N/A
7	1-5	0.0342** (0.0110)	17.69	32.25%	11,022	N/A	N/A	N/A	N/A	N/A	N/A
8	4-5	0.0603* (0.0294)	12.679	32.00%	5,443	6-8	0.0145 (0.0257)	10.59	6.00%	7,436	0.0458 (0.039)

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A9. Effects of 2023 Summer School on Fall MAP Growth Test by Grade Bands, Reading

	(A) Rising Grades 1-5					(B) Rising Grades 6-8					Difference (A) - (B)
District	Grades in Sample	Any SS 23 (SE)	Average Dosage (Hours)	% Treated	N	Grades in Sample	Any SS 23 (SE)	Average Dosage (Hours)	% Treated	N	Any SS 23 (SE)
1	N/A	N/A	N/A	N/A	N/A	7-8	0.0442 (0.0649)	24.46	3.72%	6,208	N/A
2	1-5	-0.0238* (0.0113)	41.56	30.00%	11,584	7-8	0.0022 (0.0482)	17.26	15.99%	5,307	-0.026 (0.0495)
3	4	-0.0294 (0.0319)	25.93	14.77%	4,287	6-8	0.0184 (0.0250)	17.45	7.64%	12,250	-0.0478 (0.0405)
4	2-5	-0.0389** (0.0127)	19.27	11.44%	16,279	6-8	-0.0003 (0.0156)	19.16	10.49%	15,641	-0.0386 (0.0201)
5	4-5	-0.0463 (0.036)	29.33	18.05%	2,227	6-8	0.0068 (.0246)	26.92	28.14%	3,319	-0.0531 (0.0436)
6	3-5	-0.0203 (0.0279)	7.11	6.54%	3,106	6-8	-0.0278 (0.038)	N/A	4.72%	2,585	0.0075 (0.0471)
7	1-5	0.0113 (0.0124)	17.42	32.04%	10,325	N/A	N/A	N/A	N/A	N/A	N/A
8	4-5	0.0786* (0.0333)	18.272	31.00%	5,520	6-8	-0.0114 (0.0366)	10.585	6.00%	7,517	0.0900 (0.0495)

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A10. District 1 Subgroup Analysis Results

	Math (Rising Grades 7-8)				Reading (Rising Grades 7-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.090	24.5	3.7%	6,214	0.044	24.5	3.7%	6,208
	(0.070)				(0.065)			
Below 50th Percentile	0.075	25.0	8.7%	1,957	0.072	25.0	8.7%	1,952
	(0.086)				(-0.073)			
Above 50th Percentile	0.315**	23.4	0.5%	3,295	0.224***	23.4	0.5%	3,287
	(0.118)				(-0.059)			
Not FRPL	–	–	–	–	–	–	–	–
FRPL	–	–	–	–	–	–	–	–
Not Special Education	0.081	24.6	3.5%	5,823	0.051	24.6	3.5%	5,820
	(0.063)				(0.062)			
Special Education	0.437***	23.3	7.2%	391	0.539	23.3	7.2%	388
	(0.108)				(0.296)			
Not ELL	0.094	24.8	3.7%	5,728	0.029	24.8	3.7%	5,721
	(0.074)				(-0.076)			
ELL	0.148	20.2	3.9%	486	0.247	20.2	3.9%	487
	(0.187)				(0.159)			
Female	0.065	24.2	3.6%	2,993	-0.101	24.2	3.6%	2,994
	(0.056)				(0.110)			
Male	0.092	24.7	3.9%	3,221	0.090	24.7	3.9%	3,214
	(0.114)				(0.103)			
Black	0.086	24.9	8.8%	1,758	-0.001	24.9	8.8%	1,753
	(0.080)				(0.082)			
Hispanic	0.243**	24.0	2.6%	772	0.142	24.0	2.6%	771
	(0.090)				(0.175)			
Asian, Multi-racial, NHPI, Native	0.092	23.3	3.6%	308	-0.258	23.3	3.6%	305
	(0.126)				(0.147)			
White	0.101	23.5	1.4%	3,376	0.049	23.5	1.4%	3,379
	(0.104)				(0.089)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A11. District 2 Subgroup Analysis Results

	Math (Rising Grades 1-8)				Reading (Rising Grades 1-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.005	16.8	21.7%	19,928	-0.024*	29.4	21.7%	19,928
	(-0.010)				(-0.011)			
Below 50th Percentile	0.033	14.3	20.3%	18,209	0.022	41.3	22.5%	15,890
	(0.033)				(0.039)			
Above 50th Percentile	-0.243	10.5	3.3%	821	0.219	40.3	74.7%	371
	(0.167)				(0.119)			
Not FRPL	0.055	14.0	20.4%	4,723	0.025	41.5	23.8%	3,899
	(0.059)				(0.049)			
FRPL	(0.030)	14.0	21.0%	14,307	0.021	41.2	23.7%	12,362
					(0.045)			
Not Special Education	0.032	14.0	21.1%	16,009	0.012	41.4	24.3%	13,567
	(0.040)				(0.048)			
Special Education	0.000	14.0	19.4%	3,021	0.088	40.3	20.9%	2,694
	(0.041)				(0.102)			
Not ELL	0.007	14.8	17.3%	10,303	0.029	41.6	18.2%	9,226
	(0.042)				(0.051)			
ELL	0.076*	13.4	25.1%	8,727	0.017	40.9	31.0%	7,035
	(0.031)				(0.034)			
Female	0.071	13.8	21.2%	9,362	-0.013	41.5	23.9%	7,965
	(0.044)				(0.050)			
Male	0.005	14.3	20.6%	9,668	0.054	41.0	23.5%	8,296
	(0.035)				(0.037)			
Black	–	–	–	–	–	–	–	–
Hispanic	0.032	14.1	21.0%	18,287	0.025	41.2	23.9%	15,583
	(0.034)				(0.040)			
Asian, Multi-racial, NHPI, Native	–	–	–	–	–	–	–	–
White	–	–	–	–	–	–	–	–

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A12. District 3 Subgroup Analysis Results

	Math (Rising Grades 1-4 and 6-8)				Reading (Rising Grades 4 and 6-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.034**	21.0	10.1%	28,774	0.003	20.9	9.5%	16,537
	(0.016)				(0.020)			
Below 50th Percentile	0.031*	20.7	17.3%	14,966	0.000	21.0	16.3%	8,727
	(0.017)				(0.023)			
Above 50th Percentile	0.019	23.3	2.3%	13,808	0.020	20.1	1.9%	7,810
	(0.043)				(0.046)			
Not FRPL	–	–	–	–	–	–	–	–
FRPL	–	–	–	–	–	–	–	–
Not Special Education	0.027	21.0	8.9%	25,417	-0.013	20.8	8.2%	14,577
	(0.018)				(0.023)			
Special Education	0.037	20.8	19.3%	3,357	0.053	21.2	19.0%	1,960
	(0.037)				(0.052)			
Not ELL	0.034*	21.0	9.3%	26,287	0.005	20.6	8.6%	14,598
	(0.018)				(0.023)			
ELL	0.042	20.9	18.3%	2,487	-0.019	22.0	16.6%	1,939
	(0.040)				(0.048)			
Female	-0.015	20.7	9.9%	14,149	0.007	20.7	9.2%	8,101
	(0.023)				(0.028)			
Male	0.069***	21.2	10.4%	14,625	-0.006	21.0	9.8%	8,436
	(0.021)				(0.028)			
Black	0.052**	19.9	13.0%	11,855	0.022	20.4	12.8%	6,795
	(0.021)				(0.025)			
Hispanic	0.009	22.1	14.0%	5,400	0.012	21.5	12.0%	3,165
	(0.034)				(0.040)			
Asian, Multi-racial, NHPI, Native	-0.020	19.4	10.0%	1,617	–	–	–	–
	(0.085)							
White	0.019	23.1	3.8%	7,841	-0.036	21.8	3.6%	4,413
	(0.055)				(0.066)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A13. District 4 Subgroup Analysis Results

	Math (Rising Grades 2-8)				Reading (Rising Grades 2-8)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.022**	19.8	12.2%	40,660	-0.021*	19.2	11.0%	31,920
	(0.007)				(0.010)			
Below 50th Percentile	0.017	19.9	16.1%	21,496	-0.021	19.5	17.3%	15,477
	(0.009)				(0.013)			
Above 50th Percentile	0.030**	19.4	7.7%	18,942	-0.007	18.2	4.9%	16,269
	(0.012)				(0.016)			
Not FRPL	0.021	18.1	8.3%	7,219	-0.011	17.5	7.4%	7,121
	(0.022)				(0.027)			
FRPL	0.026**	20.0	13.0%	33,441	-0.018	19.5	12.0%	24,799
	(0.008)				(0.011)			
Not Special Education	0.021**	19.8	11.9%	38,106	-0.027	19.2	10.7%	29,504
	(0.007)				(0.010)			
Special Education	0.038	19.3	15.7%	2,554	0.031	19.0	14.0%	2,416
	(0.031)				(0.043)			
Not ELL	0.007	18.7	10.8%	18,535	-0.030*	18.7	10.2%	19,610
	(0.011)				(0.014)			
ELL	0.037**	20.5	13.4%	22,125	-0.003	20.0	12.3%	12,310
	(0.009)				(0.016)			
Female	0.025*	19.6	12.2%	20,102	-0.018	19.0	10.9%	15,714
	(0.010)				(0.015)			
Male	0.022*	19.9	12.1%	20,558	-0.029*	19.5	11.0%	16,206
	(0.010)				(0.014)			
Black	0.025	19.2	13.7%	5,808	-0.006	19.2	13.1%	6,202
	(0.021)				(0.024)			
Hispanic	0.026**	20.0	12.6%	30,555	-0.006	19.4	11.4%	21,167
	(0.008)				(0.012)			
Asian, Multi-racial, NHPI, Native	–	–	–	–	–	–	–	–
White	-0.023	16.6	5.1%	2,981	-0.135*	16.2	4.7%	3,164
	(0.051)				(0.064)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A14. District 5 Subgroup Analysis Results

	Math (Rising Grades 1-8)				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	0.012	21.4	23.1%	8,920	-0.011	27.6	24.1%	5,546
	(0.013)				(0.020)			
Below 50th Percentile	0.001	21.1	25.1%	6,333	-0.015	27.5	26.4%	3,657
	(0.016)				(0.025)			
Above 50th Percentile	0.045	22.3	18.3%	2,587	0.015	28.1	19.5%	1,889
	(0.028)				(0.039)			
Not FRPL	0.051	22.6	21.8%	1,489	–	–	–	–
	(0.042)							
FRPL	0.014	21.2	23.4%	7,431	-0.004	27.2	24.2%	4,681
	(0.014)				(0.023)			
Not Special Education	0.007	21.5	22.9%	7,095	-0.023	27.9	24.2%	4,584
	(0.014)				(0.022)			
Special Education	-0.036	21.0	23.8%	1,825	0.010	27.7	24.3%	4,399
	(0.043)				(0.051)			
Not ELL	0.009	21.5	23.8%	7,233	0.006	27.9	24.2%	4,584
	(0.014)				(0.023)			
ELL	0.026	21.2	20.3%	1,687	–	–	–	–
	(0.036)							
Female	0.013	21.8	22.9%	4,389	-0.024	27.7	23.9%	2,690
	(0.018)				(0.028)			
Male	0.015	21.1	23.4%	4,531	-0.030	27.6	24.2%	2,856
	(0.021)				(0.029)			
Black	0.017	22.1	29.4%	4,111	-0.004	28.1	30.4%	2,629
	(0.019)				(0.027)			
Hispanic	0.018	19.5	17.5%	1,390	–	–	–	–
	(0.047)							
Asian, Multi-racial, NHPI, Native	0.015	21.6	19.6%	1,601	–	–	–	–
	(0.044)							
White	-0.035	19.9	16.4%	1,818	-0.042	27.1	15.7%	1,109
	(0.036)				(0.077)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. Estimates for the FRPL and white subgroups in math, as well as for the male subgroup in reading, are based on slightly simplified specifications with linear baseline (Spring 2023) MAP test and without Fall 2022 test. *$p<0.05$, **$p<0.01$

Appendix Table A15. District 6 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	0.062	–	6.38%	4,767	-0.023	–	5.71%	5,691
	(0.025)				(0.022)			
Below 50th Percentile	0.067	–	8.37%	2,677	-0.037	–	7.61%	3,087
	(0.034)				(0.035)			
Above 50th Percentile	–	–	–	–	–	–	–	–
Not FRPL	–	–	–	–	–	–	–	–
FRPL	–	–	–	–	–	–	–	–
Not Special Education	0.061	–	6.49%	4,409	-0.029	–	5.75%	5,252
	(0.027)				(0.023)			
Special Education	–	–	–	–	–	–	–	–
Not ELL	0.061	–	4.59%	2767	-0.039	–	4.24%	3,441
	(0.041)				(0.039)			
ELL	0.041	–	8.85%	2,000	-0.051	–	7.96%	2,250
	(0.038)				(0.042)			
Female	0.096	–	5.83%	2,317	-0.037	–	5.17%	2,746
	(0.037)				(0.037)			
Male	0.030	–	6.90%	2,450	-0.005	–	6.21%	2,945
	(0.028)				(0.040)			
Black	0.099	–	9.90%	1,111	0.004	–	9.11%	1,372
	(0.045)				(0.048)			
Hispanic	–	–	–	–	–	–	–	–
Asian, Multi-racial, NHPI, Native	–	–	–	–	–	–	–	–
White	–	–	–	–	–	–	–	–

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A16. District 7 Subgroup Analysis Results

	Math (Rising Grades 1-5)				Reading (Rising Grades 1-5)			
Subgroups	Any Participation (SE)	Avg Dosage in Hours	% Treated	N	Any Participation (SE)	Avg Dosage in Hours	% Treated	N
Overall	0.034	17.7	32.3%	11,022	0.011	17.4	32.0%	10,325
	(0.011)				(0.012)			
Below 50th Percentile	0.046	17.2	32.1%	7,454	0.008	17.1	32.3%	6,973
	(0.014)				(0.016)			
Above 50th Percentile	-0.010	18.6	32.7%	3,568	0.046	18.1	31.4%	3,352
	(0.022)				(0.021)			
Not FRPL	0.025	18.9	32.8%	6,170	0.015	18.7	32.7%	5,518
	(0.016)				(0.017)			
FRPL	0.051	16.1	31.6%	4,852	0.008	15.8	31.3%	4,807
	(0.018)				(0.019)			
Not Special Education	0.027	17.7	31.5%	9,825	0.017	17.4	31.2%	9,183
	(0.011)				(0.013)			
Special Education	0.131	17.7	38.7%	1,197	-0.056	17.4	38.6%	1,142
	(0.051)				(0.046)			
Not ELL	0.044	17.5	34.9%	7,061	0.015	17.2	34.2%	6,840
	(0.015)				(0.015)			
ELL	0.005	18.1	27.5%	3,961	0.005	17.9	27.8%	3,485
	(0.020)				(0.023)			
Female	0.038	17.9	32.6%	5,373	0.026	17.7	32.4%	4,995
	(0.015)				(0.019)			
Male	0.037	17.4	32.0%	5,649	0.008	17.1	31.7%	5,330
	(0.017)				(0.016)			
Black	0.056	16.6	35.2%	2,254	0.006	16.4	34.5%	2,264
	(0.027)				(0.026)			
Hispanic	0.030	18.1	27.8%	4,240	0.029	18.0	28.0%	3,774
	(0.017)				(0.022)			
Asian, Multi-racial, NHPI, Native	0.052	17.0	33.3%	2,074	-0.021	16.7	32.7%	3,352
	(0.031)				(0.034)			
White	0.020	18.6	36.3%	2,454	0.040	18.2	25.7%	2,258
	(0.026)				(0.026)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$

Appendix Table A17. District 8 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	0.038	12.3	17.0%	12,879	0.034	16.7	16.0%	13,037
	(0.020)				(0.025)			
Below 50th Percentile	0.0422*	12.6	26.9%	4,462	0.028	13.6	26.5%	4,523
	(0.021)				(0.025)			
Above 50th Percentile	0.112	12.2	14.0%	8,417	0.025	19.6	13.8%	8,514
	(0.100)				(0.189)			
Not FRPL	–	-	–	–	–	–	–	–	–	–	–	–	–	–
FRPL	–	-	–	–	–	–	–	–	–	–	–	–	–	–
Not Special Education	0.028	12.2	17.0%	10,636	0.034	16.6	16.8%	10,754
	(0.023)				(0.027)			
Special Education	0.050	13.1	25.4%	2,243	0.017	16.1	25.0%	2,283
	(0.039)				(0.061)			
Not ELL	0.043	12.3	15.2%	11,184	0.024	16.7	15.0%	11,329
	(0.024)				(0.033)			
ELL	0.069	12.5	39.9%	1,695	0.0926*	16.1	39.6%	1,708
	(0.041)				(0.046)			
Female	0.046	12.1	21.9%	6,256	0.028	17.4	21.5%	6,354
	(0.028)				(0.036)			
Male	0.022	12.8	15.2%	6,623	0.045	15.4	15.1%	6,683
	(0.029)				(0.034)			
Black	0.010	11.4	37.6%	857	-0.049	14.0	37.4%	862
	(0.059)				(0.082)			
Hispanic	0.099*	12.4	24.8%	2,054	0.141**	14.8	24.4%	2,087
	(0.041)				(0.048)			
Asian, Multi-racial, NHPI, Native	0.010	12.5	15.6%	1,581	0.046	15.2	15.3%	1,611
	(0.069)				(0.084)			
White	0.029	13.7	22.6%	832	0.093	16.0	22.3%	844
	(0.078)				(0.090)			

Note. Impact estimates are reported for a subgroup only when the analytic sample for the subgroup is greater than 1,000 and the number of treated students within the subgroup is greater than 100. *$p<0.05$, **$p<0.01$