Evaluating the Impact of a Summer School Program Using a Regression Discontinuity Design[1]

Sarah S. King[2]

October 2022

Abstract

The COVID-19 pandemic led to a near-universal shift to remote and virtual learning which led to many schools developing acceleration programs to help get students back on track. This study analyzes the efficacy of one such recovery effort: a summer school program in one large urban school district in Georgia which was implemented in summer 2021 and focuses on students in elementary and middle grades. The summer school program was intended to serve students who had failed courses or were performing below grade level on exams with the hope that additional instruction would help students to get caught up. I employ a regression discontinuity design and find that the program had minimal impacts on student achievement. Further, the study notes low attendance with most attendees coming from disadvantaged backgrounds. This paper also explores potential equity impacts of the program as well as policy suggestions.

[1] I am grateful to Tim Sass, Jennifer Darling-Aduana, Jonathan Smith, Thomas Goldring, Monica Mogollon-Plazas, and Daniel Lee for their invaluable comments and feedback throughout this project.
[2] Department of Economics, Andrew Young School of Policy Studies, Georgia State University. Email: sbarry8@gsu.edu

1. Introduction

During the spring semester of the 2019-20 school year, the COVID-19 pandemic forced schools across the country to switch to remote learning. For many schools, remote learning in some fashion continued into the fall 2021 semester and beyond. For many students, remote instruction is not as effective as traditional face-to-face instruction; combined with the impacts of crisis learning, we know that crisis remote learning led to lower student achievement growth which necessitated acceleration programs to help get students back on track (CREDO 2015; Ahn and McEachin 2017; Dorn et. al. 2020; Kuhfeld et. al. 2022). In this paper, I analyze the efficacy of one such recovery effort: a 2021 summer school program implemented in a large urban school district in GA. The summer school program targeted students who had failed courses or were performing below grade level on exams with the hope that additional instruction would help students to get caught up. This paper examines the effectiveness of this 2021 summer school program on student performance using a regression discontinuity design, and specifically considers students in elementary and middle grades as these are the grades levels where students were invited based on measurable criteria and consist of the majority of attendees. I find that the program had minimal impacts for students in the middle of the test score distribution. Further, I find that attendance is significantly related to student characteristics such as family income.

1.1 Learning Loss

The COVID-19 pandemic presented a significant shock to learning for students around the world. Most studies of student achievement growth since the start of the pandemic have found lower achievement growth compared to the pre-pandemic period and compared to other schooling disruptions in the past (Pier et. al., 2021; Kuhfeld et. al., 2022). Specifically, studies have reported reductions in student achievement growth in the spring of 2020 that are comparable to the amount of time that students lost in face-to-face instruction (Dorn et. al., 2020; Engzell et. al., 2021). The reduction in achievement growth may also have important equity implications as the pandemic has continued. For instance, authors have noted that non-white students may have been more likely to receive remote education for longer which may have exacerbated existing achievement gaps, and other studies have noted widening achievement gaps, especially among high- and low-income students (Dorn et. al., 2020; Lewis et. al., 2021; Bailey et. al., 2021; Pier et. al., 2021; West & Lake, 2021; EmpowerK12, 2021; Kuhfeld et. al., 2022).

Learning loss is a concern to educators even outside of the pandemic context; further, understanding learning loss in other contexts provides some insight as to why the pandemic was so impactful for students. In general, researchers have found lower achievement by students following summer break, with the decline of test scores being close to the equivalent of one month of learning (Cooper et. al., 1996; Quinn & Polikoff, 2017). Summer learning loss appears to be greater in math than in reading, though there are mixed results regarding learning loss for different grade levels (Hanover Research, 2020). In addition, researchers have noted that learning loss may be an issue of equity and that it may differ by subgroups, though there is inconclusive evidence regarding this. In particular, while it is known that achievement gaps grow during the school year, there is mixed evidence regarding whether or not they also grow during the summer, with some studies noting that income-based achievement gaps may grow for reading, others finding that race-based achievement gaps may shrink during the summer, and others finding no significant change in achievement gaps by subgroup (Atteberry & McEachin, 2021; Kuhfeld et. al., 2021). There is also a significant strand of literature regarding the origin of achievement gaps with most research suggesting that gaps are present before children even begin school, with inconclusive results on whether these gaps grow or not during schooling (von Hippel et. al., 2018; von Hippel & Hamrock, 2019). While this paper considers a program during the COVID-19 pandemic, it also has implications for summer programs in general and thus gives rise to the importance of understanding summer learning loss. In particular, this study is focused on a diverse district which allows the findings to be used by many different districts. Further, while pre-pandemic summer programs focused on students at risk of being retained, this program focused on the bottom half of students with this study looking at students in the middle of the achievement distribution which means one can draw conclusions regarding which students should be targeted by summer programs in the future. Finally, while this study is considering a summer program during the pandemic, the results from this district are likely more generalizable than results from other districts given that the district returned to face-to-face learning early in the 2020-21 school year.

While learning loss has mostly been studied within the context of summer breaks, there has also been some research on learning loss in other contexts. One such context is student absenteeism with researchers finding that reducing student absences, particularly unexcused absences, has a significant impact on student achievement, especially in math (Gottfried, 2009; Gottfried, 2010; Gottfried, 2011; Aucejo & Romano, 2016; Santibanez & Guarino, 2020). In addition, chronic absenteeism has been linked to greater student disruptions and therefore negative spillover effects for other students in the class (Lazear, 2001; Gottfried, 2019). Another context in which learning loss has been studied is weather-related disruptions. As snow days and other inclement weather days result in fewer instructional days, they may be related to reductions in student achievement. While inclement weather has been shown to reduce student achievement, there is mixed evidence whether this effect is due to school closures or to weather-related absences and disruptions in students' personal lives (Holmes, 2002; Marcotte, 2006; Goodman, 2015).

Additionally, researchers have explored the impacts of other crisis scenarios. Jaume and Willen (2019) examined the impact of teacher strikes in Argentina which led to the closing of schools on later labor market outcomes and found worse outcomes for students who were impacted by school closures, especially in earlier grades. Another external shock to education was Hurricane Katrina which impacted the gulf coast of the United States. The hurricane forced many families to evacuate which meant that many students had to change schools or go without schooling. The literature on student achievement as a result of this shock suggests that students achieved lower test scores after the storm and that there was a gap between students who were and were not displaced as a result of the storm, with mixed results on persistence of these effects (Ward et. al., 2008; Sacerdote, 2012). Thus, the literature on summer learning loss and learning loss in other contexts suggests that time in school is important for student achievement. Given that the research on the impacts of COVID-19 is still emerging, it is still likely that the continued disruptions even after the onset of the pandemic have likely reduced student outcomes and therefore led to the need for continued remediation. However, disruptions may happen at other times so understanding their impacts and the impacts of potential remediation programs may have implications beyond the pandemic.

1.2 Summer Programs

Summer school programs are often used as a means of remediation, both pre-pandemic and as a response to lower achievement that has resulted from crisis learning during the COVID-19 pandemic. Most of the literature on summer programs suggests that the programs may have modest effects on student learning, if any, with greater gains in math than in reading (Augustine et al., 2016; Lynch & Kim, 2017; Sharp, 2018; Hanover Research, 2020; Prettyman & Sass, 2021; Pyne et. al., 2021). Further, targeted programs, programs that are at least five weeks in duration, and programs with smaller class sizes have been found to yield the greatest impacts (Sharp, 2018; McCombs & Augustine, 2021; Pyne et. al., 2021). However, when compared to other remediation methods such as after-school programs, Katzir et. al. (2013) found that, while effective, summer school programs were not as effective as after-school programs or a combination of the two.

Regarding the impacts for various subgroups, studies have found inconclusive results, with mixed evidence regarding students from different income brackets, and greater impacts for male and Latinx students (Quinn & Polikoff, 2017; Pyne et. al., 2021). Finally, while many studies have reported low attendance for summer programs, attendance is greater for programs that offer enrichment activities and other camp-like activities (Augustine et al., 2016; McCombs & Augustine, 2021). This study intends to contribute to this gap by exploring the impacts of the summer program on various subgroups in an effort to understand whether the program was successful at closing achievement gaps. One important limitation of summer programs is that they often garner low attendance, especially among disadvantaged students, a concern with the current program of study as well. However, this study can contribute to the literature in another important way by continuing to explore the impacts of summer programs for various subgroups of students, while providing insight into a potential remediation strategy as we emerge from the COVID-19 era.

1.3 The Current Study

The current study examines a summer school program administered by a large urban school district in Georgia. Students were initially invited to the program if they met certain eligibility requirements such as a failing course grade or a below grade level score on a formative assessment. However, attendance was not required of invited students and students who were not initially invited were still able to opt into the program after the initial eligibility designations were made, but before the program began. The eligibility criteria regarding test scores give rise to a fuzzy regression discontinuity analysis where students who were barely eligible can be compared to students who were barely ineligible. This study expands on the current literature by exploring the effectiveness of this summer school program on student achievement for students in the middle of the test score distribution and by exploring the impact of this program on certain subgroups of students. Given that summer school programs before the pandemic traditionally targeted students at the bottom of the test score distribution, the RD approach in this context allows us to understand whether summer school has different impacts for average students rather than low-performing students. Further, this study uses proprietary and detailed administrative data which allows for a causal estimate of the efficacy of the program. Thus, the goal of this study is not only to provide guidance to the school district and to other school districts seeking to accelerate student learning but also to contribute to the larger literature on summer programs. While the program of interest was implemented during the COVID-19 pandemic as a response to lowered achievement growth, the implications of this study can generalize to other contexts to provide researchers and policymakers a better understanding of summer programs.

2. Institutional Context

2.1 The Summer Program

Following a year of virtual and hybrid learning, many school districts including the one of interest offered summer school programs to help students to catch up. The summer program of interest, implemented in the summer of 2021, was open to students in all grade levels and used certain criteria to invite students to participate in the program. The program was intended to provide additional instruction in certain subject areas based on student needs for students in elementary and middle grades.[3] Most elementary and middle school sessions, except for middle school world language instruction, were offered in a face-to-face format.[4] This paper focuses on the elementary and middle school face-to-face programs as these were based on measurable invitation criteria and constitute the majority of attendees.

The district offered sessions in both June and July with all grades having the option to participate in June sessions and all grades except for middle grades having the option of participating in a July session. Further, elementary had the option of registering for just one or both sessions. While high school students received instruction in the specific course needed, elementary and middle school students received more generalized instruction. Elementary students received instruction in both math and reading at all sessions; middle school students registered for instruction in a specific subject area, though reading, English-Language Arts, science, social studies were delivered together and some students received instruction in math along with another subject area.

While receiving academic instruction, the summer school program also provided students with additional holistic benefits. Specifically, students who attended summer school received two free meals, including breakfast and lunch, and most received free transportation to face-to-face sessions. During the pandemic, all students had been eligible for free meals, though free or reduced-price meal benefits, whether for specific students or all students, typically stop during

[3] For high school students, the program was intended for students to make up failed or incomplete courses, or for students to take additional courses for the purpose of acceleration. Most of the high school offerings were virtual.
[4] While middle school world language instruction was delivered virtually, middle school world language eligibility does not provide a measurable criterion by which an analysis can be conducted. Further, the majority of middle school attendees were invited due to test scores and thus attended the program in-person.

the summer months, thus this may have been an important benefit of participation. In addition, the district provided families with the option to pick up meals during the summer, though this option was only available at a limited number of locations, and families had to pick up and subsequently store and make food for a full week. The program had no tuition except for high school students who wanted to do acceleration coursework.

2.2 Summer School Eligibility Criteria

The school district used 12 different eligibility criteria to determine initial eligibility for summer school; these criteria varied by grade level (see table 1). The two most frequently used criteria were scores on formative assessments and course grades. To be eligible based on test scores, a student in elementary or middle school must have scored at a level deemed "below grade level" on the middle-of-year iReady formative assessment[5] (formative assessment scores were not used to determine eligibility for high school students as students did not take the iReady test). For an elementary student to be eligible based on course grades, the student must either have an "incomplete" grade on their transcript in reading or math, therefore having the opportunity to make up this course in summer school, or they must have a current grade below a 70 percent (i.e., failing) at the time of invitation. For a middle or high school student to be eligible based on course grades, the student must have a failing (lower than a 70 percent) or "incomplete" course grade on their transcript from spring 2020, fall 2020, or spring 2021, with middle school students needing two semesters of a failing grade in the course or a single semester of a failing grade in a foreign language course, therefore giving students the opportunity to make up these credits in summer school.

[5] This test was administered at the start of the spring semester of the 2020-21 school year.

There were a few other criteria for which I do not have the necessary data, including some criteria that are more subjective and therefore do not have a clean cutoff to examine. One criterion was only for students in kindergarten or first grade and was for students who opted for remote instruction during the 2020-21 school year and either had an attendance rate below 51 percent of classes or an assignment completion rate below 80 percent. While this criterion does provide clean cutoffs and I do know which students were eligible for this reason, the school district was unable to provide data as this was based on teacher reporting and actual student attendance or assignment completion were not tracked in a way that the district would have the information. Additional reasons that a student could attend summer school were by teacher recommendation, retention consideration for elementary and middle school students, or for high school students, or course acceleration. Not only was the data for these criteria not tracked, but they are measures for which there is no consistently used underlying scale and therefore cannot be included in a regression discontinuity analysis. The final criterion for summer school eligibility related to students on an adapted curriculum such as a special education program. These students could be eligible based on not meeting certain objectives, incomplete courses, and low engagement like other students, but data for these elements is not available and therefore cannot be used in this analysis. In addition to the criteria for initial invitation to the summer school program, the district opened up registration to all students after realizing that there would be a low rate of registration among those who were invited. The main analysis focuses on students who were initially invited to participate in the program.

3. Methodology

This study employs a regression discontinuity design that exploits eligibility criteria used for a 2021 summer school program in a large urban school district. Specifically, the study exploits criteria relating to student grades and student scores on a formative assessment. This analysis also explores the impact of the program on various student subgroups.

3.1 Data and Sample

This study uses proprietary, high-quality administrative data which was provided by the school district of interest. The data includes information on student characteristics, achievement including student scores on the iReady formative assessment and course grades, and summer school performance and attendance (if applicable). A total of 38,175 students were deemed eligible by the initial eligibility criteria, including test scores, remote attendance (for kindergarten and first grade), and course grades, with additional students having the option to opt into the summer program by request or teacher recommendation. Of those, 30,603 were in elementary or middle school which accounts for 53% of all elementary and middle school students in the district. A total of 8,874 students actually attended the summer school program with 6,967 (78.5%) being students who were initially deemed eligible and 1,907 (21.5%) being students who opted into the program. Of the attendees, 5,731 were in grades 1-8, accounting for 10.9% of all students in those grade levels, with 87.9% of those students being initially invited to participate. Therefore, a total of 30,909 students were deemed eligible and initially invited to the program but did not attend summer school, with 79% of those students being in grades 1-8.

The primary way that students in grades 1 through 8 were deemed as eligible for summer school was by scoring below grade level on an iReady assessment. A total of 29,469 (77.2% of those deemed eligible for any reason, and 96.8% of all in grades 1 through 8) students were deemed eligible based on iReady scores, with 5,038 (17.2%) of those students actually attending summer school. Among elementary and middle school students, the majority of attendees came from elementary grades with grades 2-5 having an average of 826 attendees each versus grades

6-8 which have an average of 537 attendees each. Each grade level had a similar number of students who were initially invited to attend.

Eligibility by course grades was the only way that high school students could be initially deemed as eligible and was a significant way by which middle school students were deemed as eligible. However, the number of high school students for whom complete data is available is very low so they are omitted from the analysis. In addition, many high school students who attended summer school did so simply to make up credits that may not be relevant for future courses or to accelerate. As for middle school students, approximately 26% of those who were initially invited were invited due to course grades, though only 3% were invited only due to course grades. I focus the analysis on eligibility by test scores though I do perform a secondary analysis which considers eligibility due to failing course grades for middle school students.

3.2 Empirical Method

This study employs a “fuzzy” regression discontinuity design (RDD) and considers student formative test scores as the underlying continuous variable. A regression discontinuity design considers observations that are near some cutoff and compares those just below and just above with the assumption that these observations should be fairly similar other than their assignment to treatment (Hahn et. al., 2001; Imbens & Lemieux, 2008; Lee & Lemieux, 2010). In this case, the cutoff is a below-grade-level score on a student’s iReady assessment, determined by a student’s grade level and subject area. A “fuzzy” RDD (FRD) is used when assignment to the treatment is not perfectly determined by the observable threshold. In this study, the “fuzzy” design is used given that the threshold of a below-grade-level score was not a perfect indicator of a student being deemed eligible due to some unreported exceptions and additional criteria. Therefore, the FRD measures the change in test scores where there is a discontinuity, or a nontrivial jump, in the probability of assignment, which in this case is eligibility for summer school based on iReady scores.

Students near the appropriate thresholds can be categorized into four categories: compliers, never-takers, always-takers, and defiers, with these categories giving rise to intent-to-treat (ITT) and treatment-on-treated (TOT) estimators (Angrist et. al., 1996, Imbens & Lemieux, 2008; Oldenburg et. al., 2016). In this case, compliers are students who participated due to being deemed eligible for the program and did not if they were not deemed eligible. Similarly, never-takers are those students who would never participate regardless of eligibility and always-takers are those who would always opt to participate regardless of eligibility. Monotonicity assumes that there are no defiers (students who would always choose the opposite action that that which is recommended, so would choose to attend if not eligible or to not attend if deemed eligible). While none of these groups can actually be identified, I acknowledge that the RDD aims to estimate the effect of eligibility for the treatment on compliers. Thus, I estimate an average treatment effect on those students near the cutoff who are categorized as eligible (our "intend-to-treat" group) and students who were not eligible and did not attend summer school (our "intend-not-to-treat" group). In constructing the two intended treatment groups, I omit students who were not deemed eligible by iReady but attended the summer school program anyway, students who were deemed eligible by multiple criteria or by both math and reading, and high school students as they did not take the iReady assessment. Using the ITT approach is important as there may be selection bias in an eligible student opting to attend summer school. However, a limitation of this approach is that the effect of the treatment may be underestimated.

As students in grades 1 through 8 could be deemed eligible for summer school based on iReady test scores, the outcomes of interest are iReady scores for the beginning of year test of the

2021-22 school year in math and reading. Therefore, the model used for both the math and reading test score analyses is as follows:

$$iReadyF22_i = \alpha + \beta iReadyW21_i + \tau SS_i + \gamma iReadyW21_i SS_i + \delta X_i + \epsilon,$$

where *iReadyF22* is a student *i*'s iReady test score at the beginning of the fall semester of the 2021-22 school year, *iReadyW21* is a student *i*'s iReady test score in the middle of the 2020-21 school year, *SS* denotes whether a student *i* was invited to summer school based on iReady scores, *X* is a vector of controls, and ϵ is an error term.

4. Results

4.1 Descriptive Analyses

Overall, attendance was low for the summer school program with only 17% of invited students actually attending and only 11% of all elementary and middle school students attending (compared to 56% of 1-8 students receiving invitations to attend). I first observe that there are a similar number of students who were invited to participate in each grade level, though students in elementary grades were about twice as likely to attend than students in middle grades (see table 2 and figure 1). I then observe differences in eligibility and attendance due to below grade level iReady scores by gender, race, free or reduced-price lunch (FRPL) status, and English-language learner (ELL) status for students in grades 1 through 8 (see table 3). Male and female students each made up about half of all participants in the summer school program (47% female, 53% male). The majority of participants were eligible for free or reduced-price meals (78%), compared to just 44% of elementary and middle school students in the district who were FRPM eligible. English learners also made up a disproportionate amount of participants with 17% of participants being English learners compared to only 7% of 1st through 8th graders in the district being English learners.

I break these statistics down further to understand if these participation rates were due to differences in invitation (i.e., existing achievement differences which led to invitation to the program) (see figure 2). While similar proportions of male (57%) and female (55%) were initially invited to the program, I saw much different proportions for other subgroups. For FRPM-eligible students, 79% were invited to participate, compared to 38% of all students who were not FRPM-eligible. Further, 88% of all English learners were invited to participate, compared to just 54% of native English speakers. I then can consider attendance conditional on invitation (see figure 3). Again, comparable proportions of male (18%) and female (17%) invitees opted to participate. As for FRPM-eligible students, 22% of invitees opted to participate, compared to only 9% of non-FRPM eligible invitees. Finally, 28% of English learners who were invited opted to participate, compared to 16% of native English speakers.

Given the observed attendance for various groups of students, I conduct first stage descriptive analyses using OLS models to better understand which students opted to attend the summer program and to check for a relationship between invitation and participation (see table 4). As expected from the summary statistics, being eligible for free or reduced-price meals increased the likelihood that a student would attend the program by approximately 14% and being an English learner increased the likelihood that a student would attend by approximately 12%. These figures remain significant even when controlling for a student's middle-of-year iReady score and invitation. However, gender is not a significant predictor of whether a student would attend the program. Further, I consider the impact of winter test scores on the likelihood of attendance and find that lower test scores significantly increased the likelihood of attending with 10 scale points lower leading to a 1% increase in the likelihood of attendance, conditional on invitation and other observed characteristics. Finally, I consider the first stage of whether invitation predicts attendance and find that receiving an invitation to the summer program increases the likelihood that a student attends the program by 5%. Further, given that the regression discontinuity design considers a subset of students near the threshold of invitation, I run the first stage with a bandwidth of 20 and find similar levels of significance.

4.2 Regression Discontinuity Analysis

The main specification considers all students in grades 1 through 8 with 58% of all students in these grades being invited to participate and 96.8% of those students being invited due to having a below grade level score in math or reading. I conduct this analysis for all grade levels pooled together using recentered test scores which tell us how far a student's score is from their grade and subject combination. For each case, I use the student's lowest score in winter 2021, relative to the threshold, as the underlying variable given that being below grade level in just one subject would yield a student eligible. I repeat each analysis for the outcome variables of both math scores in fall 2021 and reading scores in fall 2021. The main analysis used a bandwidth that was determined to be optimal by the RDRobust code developed by Calonico et. al. (2017) in order to minimize bias. This optimal bandwidth was determined to be 28.437 scale points for math and 23.767 scale points for reading.

For each analysis, I do not find impacts that are significantly different from zero (see tables 5-6 and figures 4-5). I find that students who were invited to the program performed 0.25 scale points higher than non-invited counterparts in math and 1.41 scale points higher in reading, though neither of these estimates are statistically significant. Further, when controlling for prior test scores from the fall of the 2020-21 school year, I continue to find insignificant results with attendees scoring 0.56 scale points higher in math and 1.19 scale points higher in reading than non-invitees. Finally, I separate elementary and middle school students to see if the program impacted students at the two levels differently. Invited students in elementary grades (1-5) scored 0.98 scale points higher in math and 1.85 scale points higher in reading than non-invitees in elementary grades while invited students in middle grades (6-8) scored 1.91 scale points lower in math and 2.28 scale points lower in reading than non-invitees, though none of these figures are significantly different from zero.

In addition to the main analysis, I am interested in impacts on various subgroups. Therefore, I also conduct analyses which specifically include only Free or Reduced Price Meal (or non-FRPM) students, only English Language Learner (or non-ELL) students, only male or female, and only students in certain regions of the district (see table 7). I am specifically interested in this to understand whether the program was effective at closing achievement gaps that have widened as a result of the pandemic. Further, I conduct separate analyses for elementary and middle school students as prior studies have found differential impacts for the two levels.

4.3 Alternative Specifications and Validity Testing

In addition to the main specification, I consider alternative specifications to test for robustness and validity of the main result. First, I consider alternative bandwidths. The bandwidth is the specific area around the threshold which is used for the regression discontinuity analysis. For this test, I alter the bandwidths to see if a larger or smaller range of values impacts our results. I first consider larger bandwidths of 35 scale points for math and 30 scale points for reading. I then consider smaller bandwidths of 20 scale points for math and 15 scale points for reading. In each case, our results do not differ significantly from the original results (see table 8).

In addition to the bandwidth tests, I conduct placebo tests. As the threshold in this analysis is based on below iReady scores for grade and subject area combinations, I expect that any impact of the program would be noticed at that threshold. Further, I expect to not see jumps at other points in the distribution as there should be no other variations in assignment to treatment. Therefore, placebo testing ensures that the only source of variation is at the expected cutoff and thus considers the impact at placebo points. As the threshold used is at 0, I consider placebo cutoffs at -30 and +30 for both math and reading. I notice that none of the placebo thresholds yield a coefficient which is significantly different from zero which suggests that there is no other discontinuity in our distribution (see table 9).

I also conduct a test for observed covariates. A regression discontinuity design assumes that observations are similar near the cutoff as assignment to treatment should be random at the threshold. This means that controlling for observed characteristics should not impact the results of the analysis. Therefore, the test for observed covariates introduces observed characteristics including gender, FRPM status, English Learner status, and grade level to see if introducing these changes the results. I find that the introduction of observed covariates does not change our initial findings for either math or reading (see table 10). In addition to testing for observed covariates, I also test for balance of covariates. As I assume randomness at the threshold, I want to ensure that observed covariates are balanced at the cutoff. I conduct this test by running the RDD analysis using each covariate as the dependent variable. The covariates used in this test include FRPM eligibility, gender, English language learner status, race, and ethnicity. I find in each case that there is no difference in the share of students belonging to certain subgroups on each side of the invitation threshold (see table 11). Therefore, I believe that we have a balanced sample based on observed characteristics.

The final alternative specification that I use to test for validity of the model is a polynomial ordering test. For the main analysis, I consider linear trends, or a polynomial of order

1, and an order of 2 for bias correction. For this test, I consider alternative polynomial orderings to ensure that our results are not sensitive to the order. I observe that none of the coefficients generated by this test are different from those found in the main results (see table 12). In addition to alternative specifications to test for validity. I use a local polynomial density estimation to test for manipulation at the threshold, based on Cattaneo et.al. (2020). The manipulation test simply is checking for whether units assigned to treatment can manipulate their assignment by checking if there is a disproportionate number of students just above and below the threshold. In this case, I know intuitively that students could not alter whether they were above or below grade level on their iReady assessment and that they had no incentive to manipulate results given that the summer program had not been announced at the time of the assessment, invitation was not binding, and the test itself was not a high-states exam. However, I confirm with the test that there was no manipulation at the threshold as the density of students falling just above or just below is similar. While students could opt to attend or not attend the program, invitation to the program due to iReady scores appears to have not been impacted by student desires (see table 13). I also notice this visually with the distribution of math and reading scale scores being fairly normal with most of the density near the invitation threshold (see figures 6-8).

4.4 Regression Discontinuity for Eligibility by Course Grades

While the main analysis of this study considers eligibility due to below iReady test scores, I also observe that many middle school students were deemed eligible due to failing course grades. Thus, I conduct a similar regression using students who were eligible due to failing course grades versus students who were not eligible at all and did not attend the program. I note that only 26 percent of invited middle school students were invited due to failing course grades, and only 4 percent of middle school students were invited *only* due to failing course grades and not some additional reason (see table 15). Thus, while I do present the results of this analysis, I acknowledge that the sample size for students invited to participate is very low, and that a very small number of students are added to our analysis by examining eligibility due to course grades. Our main specification excludes students who failed a course but were not invited and students who attended the program but did not fail a course.

I find that students who were invited to participate in the program due to failing a math course scored 14 scale points higher in math which is statistically significant. This figure grows to 16 scale points when including students who were previously excluded. However, I do not find impacts that are significantly different from zero when considering reading scores for students who failed an English-Language Arts course with invitees scoring 5 points lower than non-invitees and 5 points higher when considering the full sample (see table 16).

Given these results, I also consider the validity of this analysis. While the results are interesting, particularly for math, when I conduct the test for observed covariates, the aforementioned findings do not hold. That is, when I include observed covariates of gender, free or reduced-price meal status, and English-learner status into the model, the impacts for math become insignificantly different from zero. This suggests that, while there may be a correlation between invitation due to math grades and achievement, I cannot rule out the possibility that the impact I observed is in fact due to existing differences between the two groups of students.

4.5 Correlational Analyses

While the main analysis did not indicate that the program was effective for students near the threshold, I conduct correlational tests to understand if the program was related to student achievement at all. Interestingly, conditional on test scores on the winter assessments during the 2020-21 school year, invited students who attended scored 2.5 scale points lower in math and 6 scale points lower in reading than invitees who did not attend, with both figures being statistically significant. In addition, I evaluate whether attending the summer program for longer had any relationship to achievement as students in elementary grades could opt to attend just one or both sessions. I find that, conditional on winter test scores, students who attended both sessions did not perform significantly differently from students who only attended one session.

In a further attempt to understand the impact that this program may have had on widening achievement gaps, I perform correlational analyses on test scores at the beginning of the fall semester of the 2020-21 school year, conditional on winter of the 2020-21 school year, for students who attended the summer program by various subgroups (see table 17). By conditioning on scores when the students were deemed eligible for summer school, our results tell us how achievement gaps changed for students based on whether they opted to attend or not. I also acknowledge that, given the literature on summer learning loss, widening of achievement gaps is expected so I are interested in how achievement gaps widened between attendees and non-attendees.

First, while achievement gaps for male and female students, where female students perform slightly better than male students, did not change after the summer program in math, I do observe that the gap grew about one scale point for attendees compared to non-attendees. Further, I observe that the gap grew for all students in reading but that the gap grew by about 6 scale points more for invited students who attended the program compared to those who did not attend. This suggests that the summer program is related to widened gender achievement gaps though it is uncertain whether that is due to differences in the students who opted to attend or the program itself. Next, I consider the gap between English learners and English speakers where English speakers perform better than English learners. The gap in both math and reading did not change significantly after the summer program for those students who attended, though it grew significantly for students who did not attend the program but were invited. Further, the gap grew about 3 scale points more in both math and reading for students who did not attend compared to students who did attend which suggests that the program may have mitigated the impacts of summer learning loss for those who attended.

Finally, I consider income-based achievement gaps by looking at both free or reduced-price meal eligibility and whether a student lives in the more or less affluent portion of the district. For both cases, the existing achievement gap had students from more affluent backgrounds performing better than students from less affluent backgrounds so FRPM status and location both help us to proxy student socioeconomic status. Both measures suggest that achievement gaps grew in both math and reading after the summer program, which is expected due to previous literature on summer learning loss. However, I are more interested in how achievement gaps grew for students who participated in the program versus those who did not. When considering FRPM eligibility status, I observe that the achievement gap grew by about 2 scale points more in math and 6 scale points more in reading for invited students who did not attend the program compared to invited attendees. As for location, I observe that the gap grew about 3 scale points more in math and 7 scale points more in reading for invited non-attendees compared to invited attendees. This suggests that participation in the summer program is related to less growth in achievement gaps, though it is not certain whether this is due to unobserved differences in students who opted to attend or the program itself.

5. Discussion

In this paper, I evaluate the impact of a summer school program in 2021 on student achievement in the following school year. This program was developed in response to lowered student achievement growth during the COVID-19 pandemic. I found the program to have minimal impacts on student test scores in both math and reading; I found this result to hold through multiple specifications and for all grade levels. While these results align with the literature regarding effectiveness in reading achievement, I do note that many studies do find impacts in math which are not observed in this study. These findings are likely due to the nature of the program in that it focused on generalized instruction and yielded low attendance.

Further, I observe that the vast majority of students who took part in the program were deemed as eligible for free or reduced-price meals. I also note existing achievement gaps as most FRPM-eligible students and most English Language Learning students in the district met the initial criteria for invitation to the program. Therefore, I explore impacts of the program on various groups of students but do not find the program to be effective in closing achievement gaps even among students who attended and found students from more advantaged backgrounds to fare better following the program. Therefore, our results reveal important implications not only for understanding summer school programs in general but also for how they can impact students from various backgrounds.

5.1 Limitations and Concerns

One main concern of this study is the lack of participation in the summer school program. As attendance was only recommended but not required of eligible students, many opted to not participate in the program. This leads to a relatively low sample size to work with, though the district is large enough that there is still a significant amount of students in the sample. Additionally, while all forms of non-compliance are important to consider, the number of students who attended but were not initially deemed as eligible is very low making the primary concern the never-takers. A second concern is a lack of density around the cutoffs for some grade and subject combinations which means that the RD analysis is not be as clean as hoped. While the iReady assessments are normed to have a normal distribution across all test takers, the grade and subject combinations for this district show that there is a decline in students receiving a certain score right around the below-grade-level threshold which shows in the breakdown of students who were eligible for and attended summer school.

A final concern of this study is that there were multiple eligibility criteria and that underlying data for some of these was not available. While I attempt to address this by restricting our sample to subgroups of students eligible by only certain criteria, we are left without the full picture of the impact of summer school on some students, particularly those on an adapted curriculum and students who attended based on a teacher recommendation but did not meet one of the formal eligibility criteria. I also acknowledge that using a regression discontinuity design does not yield generalizable results but rather suggests a certain effect for those near the cutoff (known as the "local area treatment effect" or LATE); that is, I am unable to conclude whether this summer school program was effective for all students but I am able to conclude whether it was effective for those students who were just below grade level.

5.2 Policy Implications

Given that this program had minimal impacts on student achievement, I argue that districts that wish to implement summer programs in the future consider a few elements for their programs. First, we know from the literature that programs that better target student needs are more effective at boosting student achievement. This program did not specifically target student needs in certain areas and therefore believe that districts should consider this when planning programs. Further, the literature is mixed on the efficacy of summer programs alone. Therefore, I believe that districts should consider alternative or complementary programs such as after school programs for improving student outcomes.

While most summer programs aim to target academic outcomes for students, I argue that summer programs can play alternative roles in student success. First, many authors have noted that summer programs can improve student social-emotional outcomes. Therefore, it may be worthwhile for districts to focus summer programs on non-academic outcomes. Further, as our program mainly served students from disadvantaged backgrounds, I believe that there is a need in this district for holistic services during the summer. Specifically, as the summer program provided meals, I theorize that this benefit may have been an important factor in families deciding to take advantage of the program. Finally, as low attendance is a common problem among summer programs, districts may consider alternative ways to entice students to participate such as through full-day programs which then solve a childcare problem, one that I believe families in our district may have faced due to higher attendance among elementary-aged students. Further, programs with an element of "fun" such as camp-like activities, crafts, and games have been shown to have higher attendance.

5.3 Conclusion

I use a regression discontinuity design to analyze a summer school program implemented in 2021 in response to the COVID-19 pandemic. I find that it did not have any meaningful impact on student achievement in the school year following the program. Further, I note that the program primarily served students from disadvantaged backgrounds but that the program did little to close achievement gaps which were widened during the pandemic. finally, this paper discusses various ways for districts to utilize summer programs to better serve students such as targeting programs towards student needs, providing holistic benefits to students both within and outside of summer school, various methods for improving attendance at summer programs, and alternative remediation programs such as after school programs.

References

Ahn, J., & McEachin, A. (2017). Student Enrollment Patterns and Achievement in Ohio's Online Charter Schools. *Educational Researcher, 46*(1), 44–57. doi:10.3102/0013189X17692999

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, *91*(434), 444–455. https://doi.org/10.2307/2291629

Atteberry, A., & McEachin, A. (2021). School's Out: The Role of Summers in Understanding Achievement Disparities. *American Educational Research Journal*, *58*(2), 239–282. https://doi.org/10.3102/0002831220937285

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. https://www.rand.org/pubs/research_reports/RR1557.html

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Yeung, N. A. (2021). Achievement Gaps in the Wake of COVID-19. In *EdWorkingPapers.com*. https://www.edworkingpapers.com/ai21-346

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for Regression-discontinuity Designs. The Stata Journal: Promoting Communications on Statistics and Stata, 17(2), 372–404. https://doi.org/10.1177/1536867X1701700208

Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, *115*(531), 1449–1455. https://doi.org/10.1080/01621459.2019.1635480

Center for Research on Educational Outcomes (CREDO). (2015). *Online charter school study.* Stanford, CA: Stanford University. Retrieved from https://credo.stanford.edu/sites/g/files/sbiybj6481/f/online_charter_study_final.pdf

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of Educational Research*, *66*(3), 227–268. https://doi.org/10.3102/00346543066003227

EmpowerK12. (2021). *COVID-19's Impact on Student Achievement and Academic Growth in DC*. https://www.empowerk12.org/research-source/covid-impact-achievement-dc

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *Mind the gap: COVID-19 is widening racial disparities in learning, so students need help and a chance to catch up | McKinsey*. McKinsey & Company. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17). https://doi.org/10.1073/pnas.2022376118

Goodman, J. (2015). In defense of snow days: Students who stay home when school is in session are a much larger problem. *Education Next*, *15*(3), 64–70.

Gottfried, M. A. (2009). Excused Versus Unexcused: How Student Absences in Elementary School Affect Academic Achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 392–415. https://doi.org/10.3102/0162373709342467

Gottfried, M. A. (2010). Evaluating the Relationship Between Student Attendance and Achievement in Urban Elementary and Middle Schools: An Instrumental Variables

Approach. *American Educational Research Journal*, *47*(2), 434–465. https://doi.org/10.3102/0002831209350494

Gottfried, M. A. (2011). The Detrimental Effects of Missing School: Evidence from Urban Siblings. *American Journal of Education*, *117*(2), 147–182. https://doi.org/10.1086/657886

Gottfried, M. A. (2019). Chronic Absenteeism in the Classroom Context: Effects on Achievement. *Urban Education*, *54*(1), 3–34. https://doi.org/10.1177/0042085915618709

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Holmes, G. M. (2002). Effect of Extreme Weather Events on Student Test Performance. *Natural Hazards Review*, *3*(3), 82. https://doi.org/10.1061/(ASCE)1527-6988(2002)3:3(82)

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.

Katzir, T., Goldberg, A., Aryeh, T. J. B., Donnelley, K., & Wolf, M. (2013). Intensity vs. Duration: Comparing the Effects of a Fluency-Based Reading Intervention Program, in After-School vs. Summer School Settings. *Journal of Education and Training Studies*, *1*(2), 61–73.

Kuhfeld, M., Condron, D. J., & Downey, D. B. (2021). When Does Inequality Grow? A Seasonal Analysis of Racial/Ethnic Disparities in Learning From Kindergarten Through Eighth Grade. *Educational Researcher*, *50*(4), 225–238. https://doi.org/10.3102/0013189X20977854

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test Score Patterns Across Three COVID-19-impacted School Years. In *EdWorkingPapers*. https://www.edworkingpapers.com/ai22-521

Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, *116*(3), 777–803. https://doi.org/10.1162/00335530152466232

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Lewis, K., Kuhfeld, M., Ruzek, E., & McEachin, A. (2021). *Learning during COVID-19: Reading and math achievement in the 2020-2021 school year*. Center for School and Student Progress. https://www.nwea.org/research/publication/learning-during-covid-19-reading-and-math-achievement-in-the-2020-2021-school-year/

Lynch, K., & Kim, J. S. (2017). Effects of a Summer Mathematics Intervention for Low-Income Children: A Randomized Experiment. *Educational Evaluation and Policy Analysis*, *39*(1), 31–53. https://doi.org/10.3102/0162373716662339

Marcotte, D. E. (2007). Schooling and test scores: A mother-natural experiment. *Economics of Education Review*, *26*(5), 629–640. https://doi.org/10.1016/j.econedurev.2006.08.001

McCombs, J., & Augustine, C. H. (2021). *Advancing Student Learning and Opportunity through Voluntary Academic Summer Learning Programs*. RAND Corporation. https://annenberg.brown.edu/recovery/EdResearch2

Oldenburg, C. E., Moscoe, E., & Bärnighausen, T. (2016). Regression Discontinuity for Causal Effect Estimation in Epidemiology. *Current Epidemiology Reports*, *3*, 233–241. https://doi.org/10.1007/s40471-016-0080-x

Pier, L., Christian, M., Tymeson, H., & Meyer, R. H. (2021). *COVID-19 Impacts on Student Learning*. https://edpolicyinca.org/publications/covid-19-impacts-student-learning

Pier, L., Hough, H. J., Christian, M., Bookman, N., Wilkenfeld, B., & Miller, R. (2021, January 25). *COVID-19 and the Educational Equity Crisis*. Policy Analysis for California Education. https://edpolicyinca.org/newsroom/covid-19-and-educational-equity-crisis

Prettyman, A., & Sass, T. (n.d.). *Impact of Virtual Summer School on Student Achievement Growth During COVID-19*. Georgia Policy Labs. Retrieved February 11, 2022, from https://gpl.gsu.edu/publications/the-impact-of-virtual-summer-school-on-student-achievement-growth-during-covid-19/

Pyne, J., Messner, E., & Dee, T. S. (2021). The Dynamic Effects of a Summer Learning Program on Behavioral Engagement in School. *Education Finance and Policy*, 1–51. https://doi.org/10.1162/edfp_a_00368

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

Sacerdote, B. (2012). When the Saints Go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita. *American Economic Journal: Applied Economics*, *4*(1), 109–135. https://doi.org/10.1257/app.4.1.109

Sharp, C. (2018). *Can Summer Schools Improve Outcomes for Disadvantaged Pupils? NFER Social Mobility Briefing* (National Foundation for Educational Research. The Mere, Upton Park, Slough, Berkshire, SL1 2DQ, UK. Tel: +44-1753-574123; Fax: +44-1753-637280; e-mail: enquiries@nfer.ac.uk; Web site: http://www.nfer.ac.uk; National Foundation for Educational Research). National Foundation for Educational Research.

https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=eric&AN=ED590446&site=eds-live&scope=site&custid=gsu1

Hanover Research (2020). *Summer Learning Loss Literature Review*. https://www.hanoverresearch.com/reports-and-briefs/summer-learning-loss-literature-review/

von Hippel, P. T., & Hamrock, C. (2019). Do Test Score Gaps Grow before, during, or between the School Years? Measurement Artifacts and What We Can Know in Spite of Them. *Sociological Science*, *6*, 43–80. https://doi.org/10.15195/v6.a3

von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, of "Are Schools the Great Equalizer?" *Sociology of Education*, *91*(4), 323–357. https://doi.org/10.1177/0038040718801760

Ward, M. E., Shelley, K., Kaase, K., & Pane, J. F. (2008). Hurricane Katrina: A Longitudinal Study of the Achievement and Behavior of Displaced Students. *Journal of Education for Students Placed at Risk (JESPAR)*, *13*(2–3), 297–317. https://doi.org/10.1080/10824660802350391

West, M. R., & Lake, R. (2021). How Much Have Students Missed Academically Because of the Pandemic? A Review of the Evidence to Date. In *Center on Reinventing Public Education*. Center on Reinventing Public Education. https://eric.ed.gov/?id=ED614130

Wong, V. C., Steiner, P. M., & Cook, T. D. (2013). Analyzing Regression-Discontinuity Designs With Multiple Assignment Variables: A Comparative Study of Four Estimation Methods. *Journal of Educational and Behavioral Statistics*, *38*(2), 107–141. https://doi.org/10.3102/1076998611432172

Table 1. Summer School Invitation Criteria

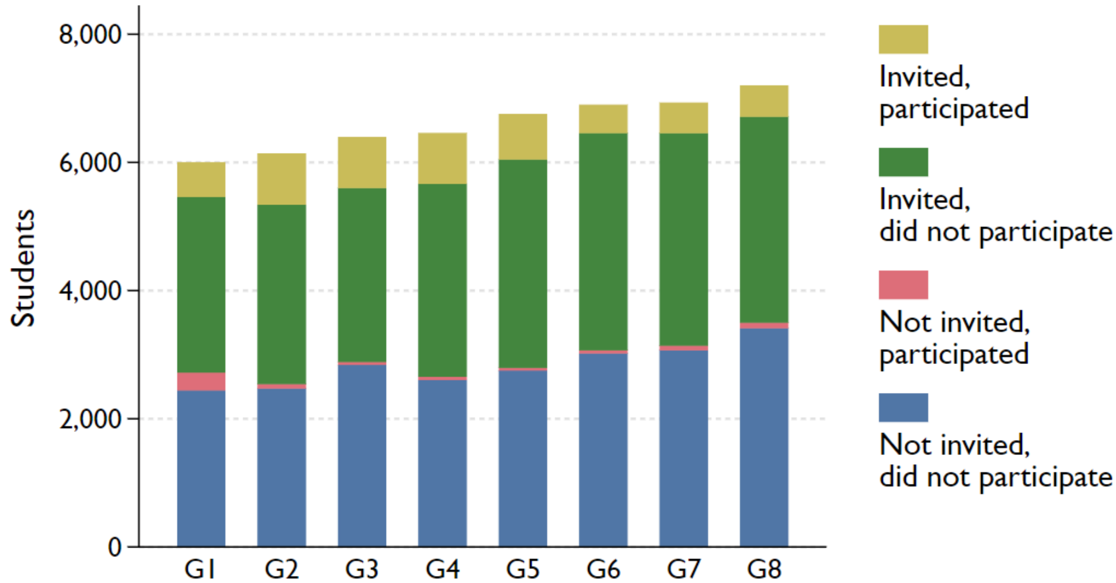
	K	1	2	3	4	5	6	7	8	High School
Remote Attendance/Completion	X	X								
Incomplete Reading Grade			X	X	X	X				
Incomplete Math Grade			X	X	X	X				
iReady, Reading		X	X	X	X	X	X	X	X	
iReady, Math		X	X	X	X	X	X	X	X	
Failed Course							X	X	X	X
Incomplete Course							X	X	X	X
World Language							X	X	X	
Acceleration										X
Teacher Recommendation	X	X	X	X	X	X	X	X	X	X
Retention Consideration	X	X	X	X	X	X	X	X	X	
Adapted Curriculum	X	X	X	X	X	X	X	X	X	X

Notes. Students could opt in for the summer program in all grade levels.

Table 2. Summer School Invitation and Participation by Grade Level

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Total
Eligible and Attended	539	801	795	791	709	440	474	489	5,038
Eligible and Not Attended	2,741	2,798	2,714	3,011	3,251	3,389	3,315	3,212	24,431
Not Eligible and Attended	276	68	47	51	42	48	74	87	693
Not Eligible and Not Attended	2,444	2,474	2,841	2,604	2,754	3,022	3,069	3,415	22,623
Total	6,000	6,141	6,379	6,457	6,756	6,899	6,932	7,203	52,785

Figure 1. Summer School Invitation and Participation by Grade Level

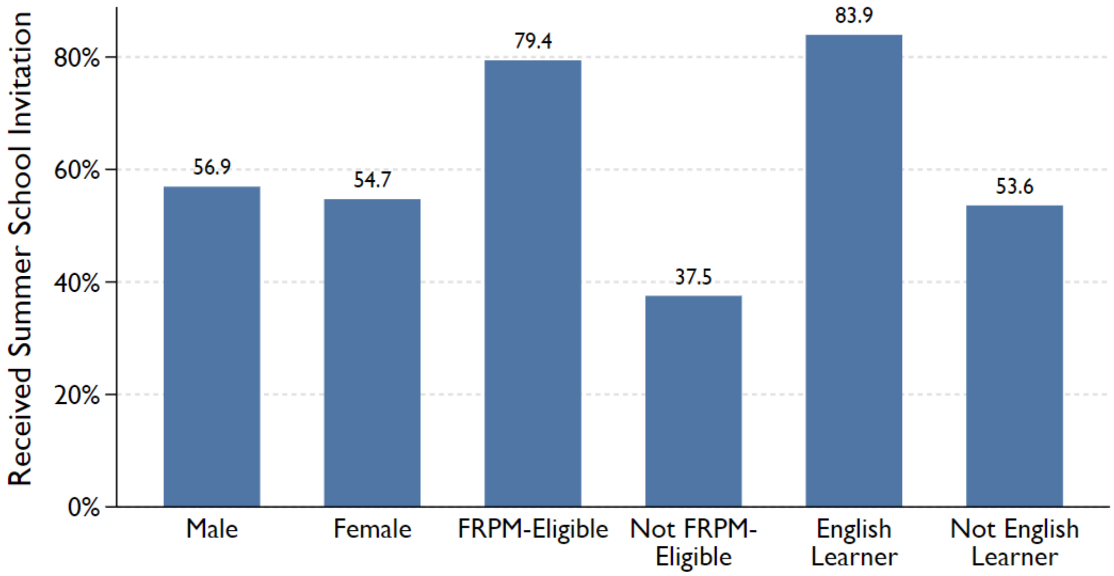


Notes. Figure shows the number of students who received an invitation due to having a below grade level score on the middle of year iReady assessment in the 2020-21 school year and their subsequent participation in summer school.

Table 3. Summer School Invitation and Participation by Subgroup

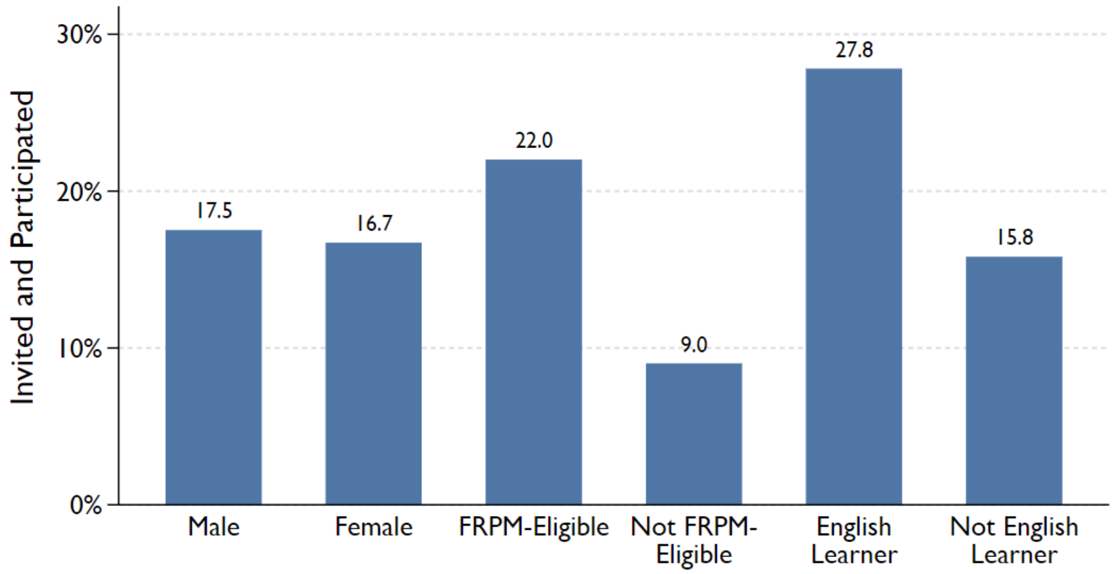
	Female	Male	Not FRPM-Eligible	FRPM-Eligible	Not ELL	ELL	Total
Eligible and Attended	2,372 (47.1%)	2,666 (52.9%)	1,001 (19.9%)	4,037 (80.1%)	4,137 (82.1%)	901 (17.9%)	5,038
Eligible and Not Attended	11,851 (48.5%)	12,580 (51.5%)	10,101 (41.3%)	14,330 (58.7%)	22,093 (90.4%)	2,338 (9.6%)	24,431
Not Eligible and Attended	338 (48.8%)	355 (51.2%)	253 (36.5%)	440 (63.5%)	617 (89.0%)	76 (11.0%)	693
Not Eligible and Not Attended	11,432 (50.5%)	11,191 (49.57%)	18,286 (80.8%)	4,337 (19.2%)	22,079 (97.60%)	544 (2.4%)	22,623
Total	25,993 (49.2%)	26,792 (50.8%)	29,641 (56.2%)	23,144 (43.8%)	48,926 (92.69%)	3,859 (7.3%)	52,785

Figure 2. Summer School Invitation by Subgroup



Notes. Percentages denote the percentage of students in each subgroup who received an invitation due to below grade level score on the middle of year iReady assessment in the 2020-21 school year.

Figure 3. Summer School Participation among Invitees



Notes. Percentages denote the percentage of students in each subgroup who were invited due to having a below grade level iReady score who opted to attend the summer program.

Table 4. First Stage Statistics

			Bandwidth	
			Full Sample	(-20, 20)
FRPM	0.1431***	0.0958***	0.0869***	0.0639***
	(0.003)	(0.003)	(0.003)	(0.004)
Female	-0.0064	-0.0006	-0.001	-0.0024
	(0.003)	(0.003)	(0.003)	(0.003)
ELL	0.1211***	0.0949***	0.0938***	0.0931***
	(0.005)	(0.006)	(0.005)	(0.010)
MOY Lowest		-0.0009***	-0.0007***	-0.0009***
		(0.000)	(.000)	(0.005)
Invitation			0.0522***	0.0281***
			(0.004)	(0.005)
N	52,785	46,404	46,404	16,876

Notes. Standard deviations in parentheses. Asterisks denote statistical significance, *** p<0.001, ** p<0.005, * p<0.01

Table 5. RDD Results for Fall 2021 Math Scores

	Coefficient	Standard Error	P > |z|
Invited	-0.250	0.599	0.677
Elementary Grades	-0.968	0.749	0.196
Middle Grades	1.910	1.428	0.181
Control for Prior (Fall 2020) Scores	-0.564	0.617	0.361
		Invited	Not Invited
N	27,552	15,926	11,626
N (within bandwidth)	14,108	6,599	7,509
N (elementary grades)	18,186	10,498	7,688
N (middle grades)	9,366	5,428	3,938

Notes. The main specification for these results uses a restricted sample of students who were invited to participate due to iReady scores and students who were not invited to participate due to iReady scores and did not participate. In addition, students who participate but were not invited are omitted.

Figure 4. RDD Analysis of Fall 2021 Math Scores

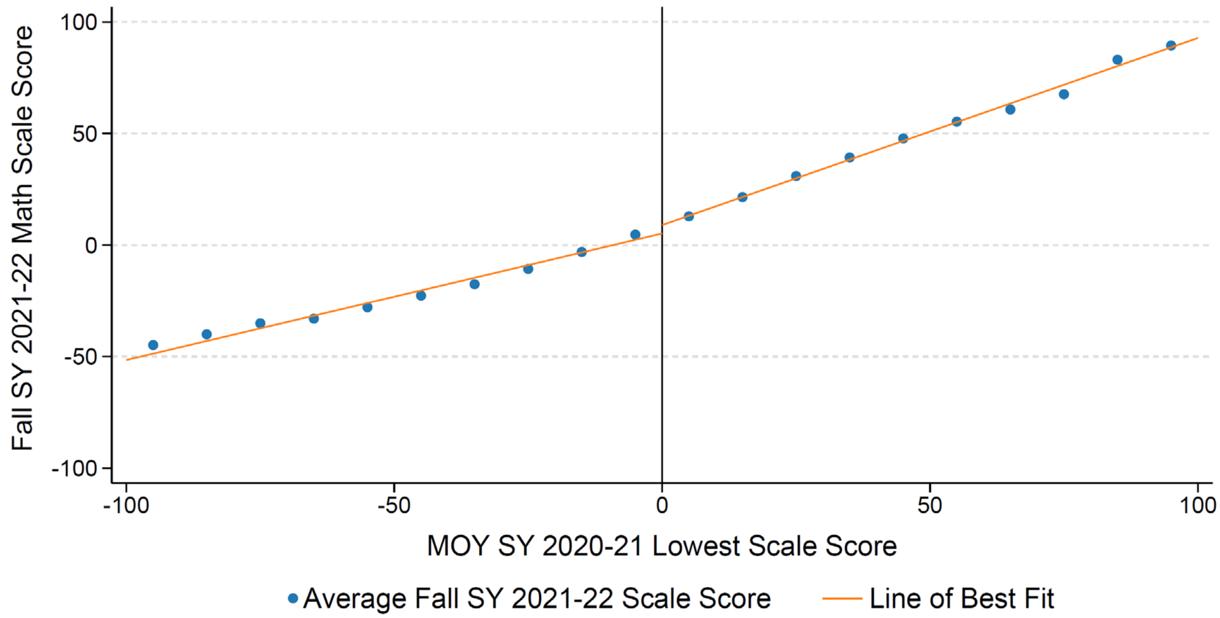


Table 6. RDD Results for Fall 2021 Reading Scores

	Coefficient	Standard Error	P > |z|
Invited	-1.405	1.334	0.292
Elementary Grades	-1.852	1.712	0.279
Middle Grades	2.282	2.030	0.261
Control for Prior (Fall 2020) Scores	-1.190	1.345	0.376
		Invited	Not Invited
N	31,163	17,370	13,793
N (within bandwidth)	13,246	6,255	6,991
N (elementary grades)	18,024	10,369	7,655
N (middle grades)	9,211	5,202	4,009

Notes. The main specification for these results uses a restricted sample of students who were invited to participate due to iReady scores and students who were not invited to participate due to iReady scores and did not participate. In addition, students who participate but were not invited are omitted.

Figure 5. RDD Analysis of Fall 2021 Reading Scores

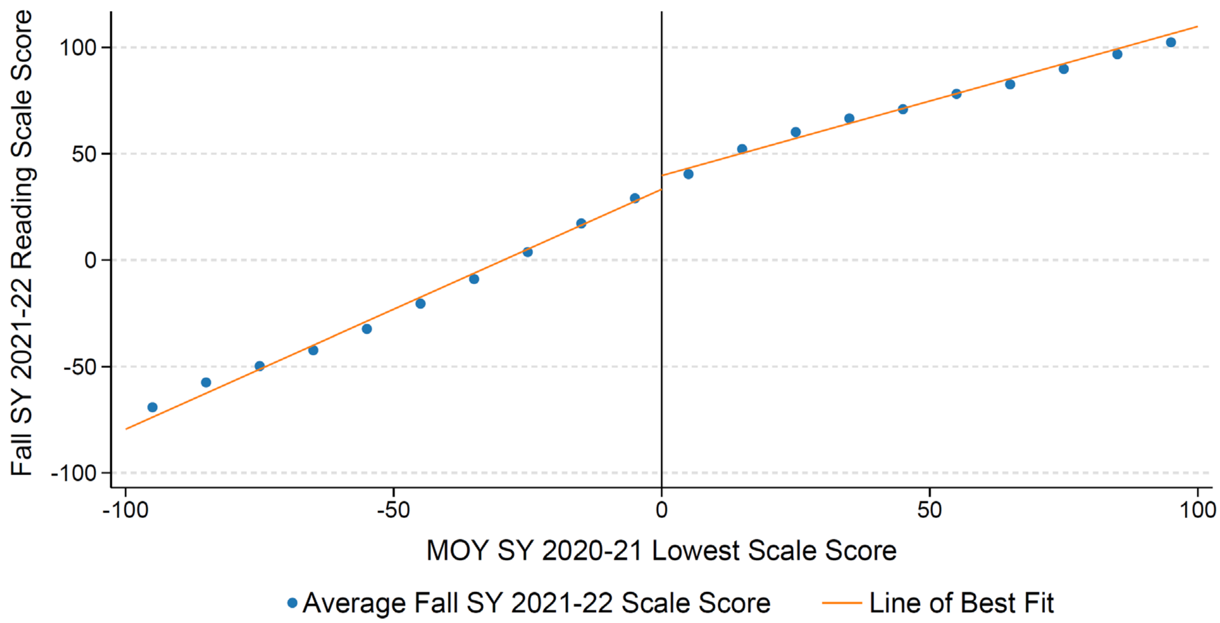


Table 7. RDD Results for Fall 2021 Scores by Subgroup

	Math	Reading
Male	0.558	-1.704
	(0.845)	(1.819)
Female	-1.471	-0.631
	(0.825)	(1.745)
FRPM-Eligible	1.772	-1.519
	(1.153)	(1.968)
Not FRPM-Eligible	-1.474	-2.129
	(0.756)	(1.704)
English Learner	-5.153	-0.197
	(5.055)	(6.125)
Not English Learner	-0.065	-1.338
	(0.600)	(1.342)
Higher-Income Region	-0.975	-1.192
	(0.811)	(1.960)
Lower-Income Region	-0.721	-3.507
	(1.544)	(3.080)

Notes. Standard errors in parentheses.

Table 8. RD Results Using Alternative Bandwidths

	Math	Reading
Original Specification	0.250	1.405
	(0.599)	(1.334)
Larger Bandwidth	-0.348	-1.071
	(0.546)	(1.193)
Smaller Bandwidth	-0.226	-1.894
	(0.705)	(1.681)

Notes. Standard errors in parentheses. The original bandwidths used were 28.4 scale points for math and 23.8 scale points for reading. The alternative bandwidths for math are 35 and 20. The alternative bandwidths for reading are 30 and 15.

Table 9. Placebo Test

	Math	Reading
Main specification (0)	0.250	1.405
	(0.599)	(1.334)
Lower Cutoff (-30)	0.825	.752
	(0.801)	(1.527)
Higher Cutoff (+30)	0.915	-1.624
	(.900)	(1.605)

Notes. Standard errors in parentheses.

Table 10. Test of Observed Covariates

Main Specification, Math	0.250	-0.511
	(0.599)	(0.576)
Main Specification, Reading	1.405	-0.476
	(1.334)	(1.171)
Controls		X

Notes. Standard errors in parentheses. Controls used include gender, FRPM eligibility, and ELL status.

Table 11. Test for Balance of Covariates

	Coefficient	Standard Error	P > |z|
FRPM Eligible	-0.012	0.015	0.433
Female	0.002	0.014	0.891
English Language Learner	0.006	0.004	0.137
Black	-0.017	0.140	0.216
Hispanic	0.007	0.009	0.481

Notes. The characteristics in the table represent dummy variables where belonging to the group yields a result of 1 and not belonging yields a result of 0.

Table 12. Polynomial Ordering Test

	Math	Reading
Main Specification	0.250	1.405
	(0.599)	(1.334)
Order 2	-0.114	-1.965
	(0.802)	(1.533)
Bias Correction Order 3	-0.223	-1.590
	(0.799)	(1.513)

Notes. Standard errors in parentheses. Main specification uses an order of 1 and bias correction of order 2.

Table 13. Density Test for Manipulation

	T	P > |T|
Equal Density	-0.467	0.641

Figure 6. Winter 2021 Math Scale Score Distribution

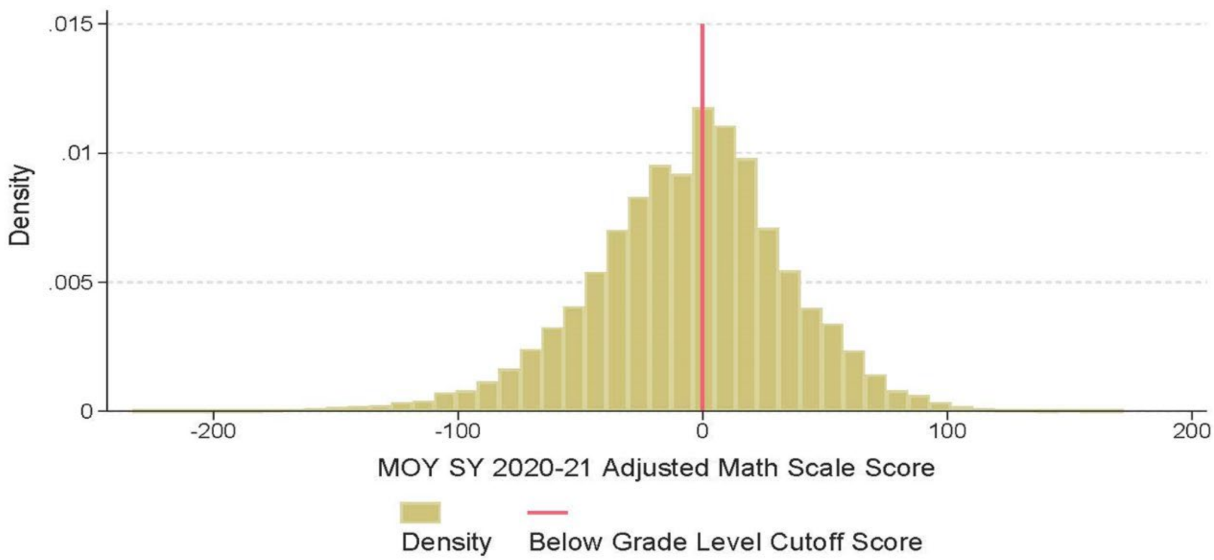


Figure 7. Winter 2021 Reading Scale Score Distribution

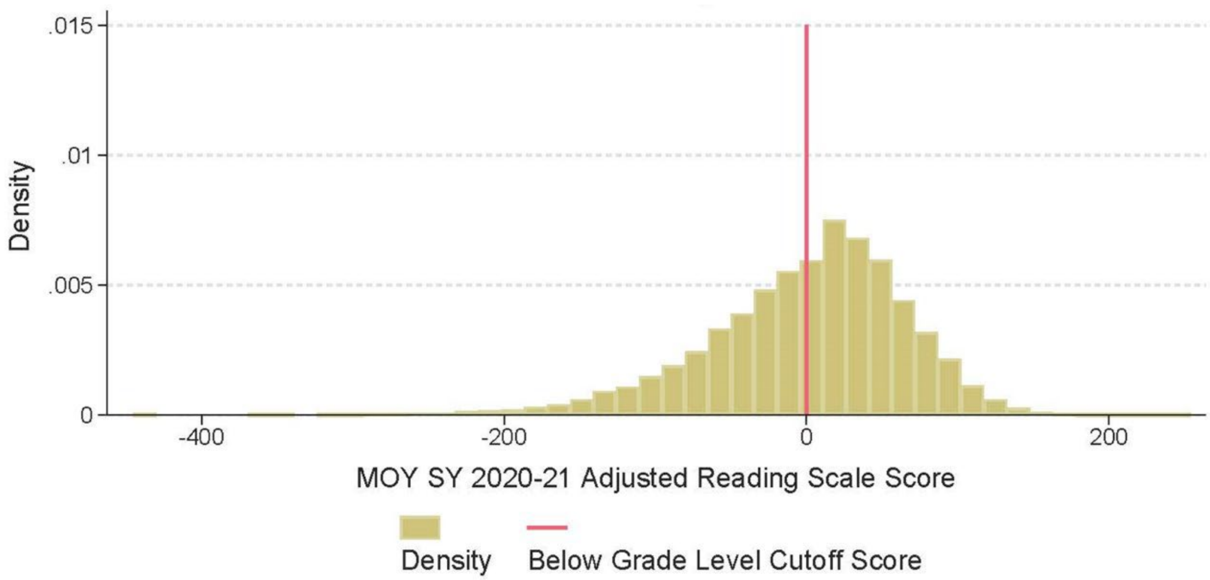


Figure 8. Density Plot

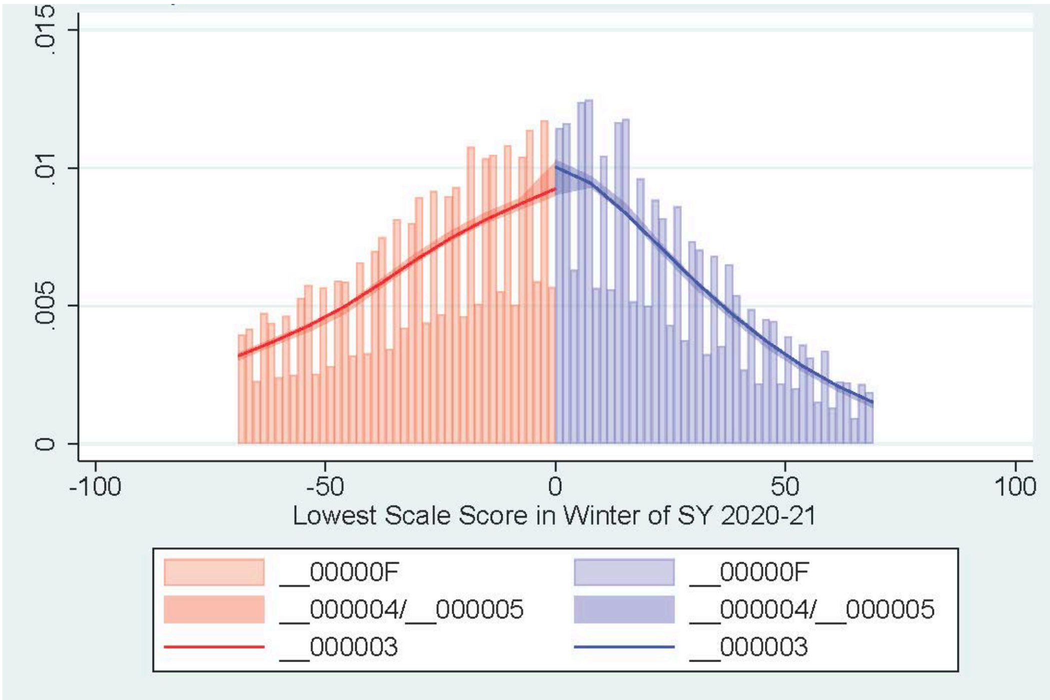


Table 14. Invitation by Criteria for Middle School Students

	Middle Grades
Only invited due to below- grade-level iReady score	9,554 (74.0%)
Only due to course failure	461 (3.6%)
Invited due to both below-grade-level iReady score and course failure	2,905 (22.5%)

Notes. Percentages based on total number of eligible students in grades 6-8.

Table 15. RDD Results of Fall Scores based on Invitation due to Grades

Math	14.068**	-8.828
	(4.651)	(4.196)
Reading/ELA	4.987	1.303
	(7.466)	(6.548)
Controls		X

Notes. Failing a math course led to an invitation to participate in math instruction, and thus should relate to future math scores. Likewise, failing an English-Language Arts (ELA) course led to an invitation to participate in reading/ELA/science/social studies instruction and thus should relate to future reading scores. Controls used include gender, FRPM eligibility status, and ELL status. Asterisks denote statistical significance, *** p<0.001, ** p<0.005, * p<0.01

Table 16. Fall 2021 Sub-group Achievement Differences Conditional on Winter 2020 Achievement Levels by Summer School Attendance

	Math			Reading		
Subgroup Comparison	Invited and Attended	Invited, Did Not Attend	Diff.	Invited and Attended	Invited, Did Not Attend	Diff.
Female vs. Male	0.337 (0.830)	-0.418 (0.369)	0.755***	6.881*** (1.337)	1.215* (0.611)	5.666***
FRPM vs. non-FRPM	-6.636*** (1.041)	-8.543*** (0.380)	1.907***	-7.788*** (1.694)	-13.425*** (0.625)	5.637***
EL vs. non-EL	1.191 (1.008)	-2.121*** (0.619)	3.312***	-1.661 (1.617)	-5.133*** (0.994)	3.472***
Lower-Income vs. Higher-Income Region	-6.927*** (0.893)	-9.646*** (0.439)	2.719***	-6.599*** (1.429)	-13.883*** (0.730)	7.284***
Attended 2 Sessions vs. 1 Session	0.368 (1.037)			-0.068 (1.663)		
Attendees vs. Non-Attendees	-2.461***			-6.207***		

Notes. Table shows impact of belonging to a given subgroup on fall 2021 test scores conditional on winter 2021 test scores for invited attendees and invited non-attendees. Parentheses denote standard errors. Asterisks denote statistical significance, *** p<0.001, ** p<0.005, * p<0.01