Gender Differences in Remote Learning amid COVID-19: Disruptive Peers and Self-Control*

Sungmee Kim†

May 2023

Abstract

The shift to remote and blended learning during pandemic-induced school closures transformed the educational landscape, altering the significance of various educational inputs and their influence on student outcomes. This paper investigates gender disparities in achievement growth during the pandemic. Utilizing Kitagawa-Blinder-Oaxaca decomposition and Two-Stage Least Squares methods, I analyze the effects of exposure to disruptive peers and gender differences in self-control on learning trajectories. Findings reveal that both factors significantly contributed to student achievement during the pandemic, with variations in self-control and peer disruption explaining a notable portion of the gender achievement gaps. Additionally, during blended learning, math achievement gaps widened for students who remained remote, particularly among low-achievers, while no significant gaps were observed for those who returned to in-person instruction. The achievement gaps were most pronounced among low-achieving students.

Keywords: Gender Achievement Gap, COVID-19, Remote Learning, Kitagawa-Blinder-Oaxaca Decomposition

JEL Codes: I21, I24, J16

---

*Acknowledgement: I would like to express my gratitude to the Georgia Policy Labs and their partner school districts for providing the data essential for this research. I also extend my thanks to Tim Sass, Daniel Kreisman, Jim Marton, Haeil Jung, Daniel Lee, and participants at the GSU Seminars, GCSU presentation, and the ASSA/AEA 2023 Annual Meeting poster session for their valuable feedback and support. This research is generously funded by the Dan E. Sweat Dissertation Fellowship at Georgia State University.

†Department of Economics, Georgia State University, seungmeeus@gmail.com

1 Introduction

The COVID-19 pandemic precipitated a global transition from traditional in-person education to remote learning, fundamentally reshaping the educational landscape and exacerbating existing challenges and disparities. This shift not only significantly impacted students' academic performance, but also raised concerns about widening preexisting achievement gaps among diverse student groups (National Center for Education Statistics, 2022; Skar et al., 2021; Bacher-Hicks et al., 2022; Aucejo et al., 2020; Copeland et al., 2021; Bailey et al., 2021; Donnelly and Patrinos, 2021; Dorn et al., 2020; Hammerstein et al., 2021; Kuhfeld et al., 2022). One of the impacts observed has been in the domain of gender-based academic achievement, as highlighted by recent report from the National Assessment of Educational Progress (NAEP). The report documented alarming trends in mathematics achievement among fourth-grade students: the gender achievement gap, previously relatively modest, has now widened to an unprecedented extent. Specifically, the gender disparity in mathematics scores has reached its highest level going back to 1990. These widening achievement gaps could leave a lasting legacy, influencing students' future outcomes and translating into larger gaps in later years if affected at younger ages (Werner and Woessmann, 2022; Autor et al., 2020).

This sudden transition to remote learning and blended learning due to the pandemic provides a unique context to examine potential factors influencing the achievement gaps. This paper, motivated by early findings from a report documenting that female students in metro-Atlanta school districts fared better than male students in reading and math during remote learning, explores trends and potential mechanisms behind the observed gender achievement gaps throughout the pandemic (Sass and Goldring, 2021). While several studies have investigated learning loss disparities among various student subgroups due to pandemic-related disruptions to learning, less is known about the underlying mechanisms driving the achievement gaps between boys and girls amidst the pandemic-induced changes in learning environments.

When the COVID-19 virus started to spread in early 2020, schools around the world responded to the situation by closing their buildings and serving students remotely for the remainder of the 2019-2020 school year, which changed the nature of student learning environment (UNESCO Institute for Statistics, 2021). Compared to a traditional face-to-face learning environment, the pandemic-induced remote learning has likely altered the way that a combination of educational inputs – such as student/family inputs, peer inputs, school/teacher inputs, et cetera – interact and affect student achievement. For instance, there had been undoubtedly fewer direct peer interactions and interactions between teachers and students, which would result in less exposure to misbehaving peers and teachers' supervision. Such changes in the relative importance of educational inputs during remote learning could be potential mechanisms that might explain widened achievement gaps across various student subgroups, given the evidence from the literature that the amount of exposure to each educational input, and the magnitude of impacts of those inputs differ across student subgroups (Krein and Beller, 1988; Dahl and Lochner, 2005; Autor et al., 2020).

While existing literature proposes several channels to elucidate the observed gender learning gaps, I focus on two potential mechanisms: classroom peer composition and students' self-control.[1] Disruptive-peer effects and innate/extrinsic self-control level vary by student gender and may potentially contribute to changes in achievement gaps between boys and girls (Zimmerman, 2003; Duckworth and Seligman, 2006; Han and Li, 2009; Ficano, 2012; Duckworth et al., 2015; Carrell et al., 2018). Moreover, there has been a growing strand of literature on the role of non-cognitive skills and peers as sources of gender gaps as well as factors determining student outcomes (Jacob, 2002; Bertrand and Pan, 2013; Nakajima et al., 2020). [Given that there was plausibly exoge-

[1] Evidence shows that K-12 gender achievement gaps can be explained by differences in behavior, cognitive and non-cognitive skills, differential impact of learning environments, and different study habits (Chen et al., 2022).

nous variation in student exposure to remote instruction in the district I study, it allows researchers to study the causal relationship between two mechanisms and the observed gender gap in academic achievement.] Based on the pandemic-induced shifts in learning environment and the evidence from the existing literature, I propose two hypotheses as potential explanations for the observed gender achievement gaps in the districts I study here: (i) remote instruction changed the nature of peer interactions and girls were less disrupted by their mis-behaving peers during remote learning after school closures in mid-March of 2020, and (ii) girls are better at self-control, which is an essential component of success in remote learning, and thus learned more than boys did when schools were closed. As the potential mechanisms may have long-term consequences for boys' and girls' learning trajectories and progressions in the future, gender differences in achievement are a matter of considerable concern (OECD, 2019).

The central questions in this paper are: (i) did remote learning after the initial school closure dampen any negative effects of having disruptive peers in classrooms? (ii) did students lacking self-control perform worse in exams after the initial remote learning period? (iii) did any observed gender differences in student outcomes during remote learning diminish for students who returned to in-person learning in Fall of SY 2020-2021? I examine these questions by utilizing administrative datasets of a metro-Atlanta school district and exploiting the variation in the intensity of classroom disruptiveness, self-control level, and the proportion of instructional remote learning days by gender. In order to explore the trend in the gender achievement gaps and estimate the change in the magnitude of impacts of the two key mechanisms across gender over the course of the pandemic, I use the Kitagawa-Blinder-Oaxaca decomposition method, where I estimate empirical models separately for female and male student groups and investigate whether changes in gender achievement gaps during the pandemic-induced remote learning stemmed from the two mechanisms of interest. To overcome potential selection bias resulting from parental choice of learning mode for their kids in the fall of 2020-2021 school year, I use Two-Stage Least Squares (2SLS hereafter) method and employ instrumental variables.

The rest of the paper is organized as follows: Section 2, comprised of two subsections, provides background. The first subsection provides information on school closures and return to in-person learning in Georgia and the particular school district I study in this paper. The second subsection documents student achievement and pre-existing gender achievement gaps in the district. A conceptual framework for the study, which is based on a traditional cumulative achievement production function, and methodology for the analyses are presented in Section 3. Section 4 provides results of the analyses, and the last section discusses the implications of the findings and concludes.

2 Background

2.1 School Closure and Return to In-Person Learning in Metro-Atlanta School Districts

Due to the impact of COVID-19, Governor Brian P. Kemp issued an Executive Order on March 14, 2020 to close all public elementary, secondary, and post-secondary schools in Georgia from March 18, 2020 through March 31, 2020 and accordingly students were offered remote learning (Lane, 2020; Sass and Goldring, 2021). Another executive order was signed on March 26 of the same year to extend the school closure through April 24, 2020 and a week later on April 1, 2020, the Governor announced all K-12 public schools would remain closed for the remainder of school year (SY) 2019-20 (Georgia Department of Education, 2020).

After the school closures in March 2020, most school districts in metro Atlanta began SY 2020-2021 with fully remote instruction but started to offer parents a choice of in-person instruction for their child at varying times in SY 2020-21. The school district I study chose a “phased” approach for returning to face-to-face instruction during the fall of SY 2020-2021. Table 1 describes phases and actual timing of return to full-time in-person instruction in the district. Each phase was implemented based on the district’s school reopening plan matrix before the district fully switched to offering full-time face-to-face learning on October 14, 2020[2]. Until the first phase (Phase I) began on September 9 of the same year, remote learning was provided to all students. Given that the school year typically begins in early August, Phase I started about a month after the school year began. During Phase I, students in Pre-K through 2 were given a voluntary opportunity to receive a 90-minute in-person instruction and support session once per week. During this phase, students in grades 3-12 were given the option to receive such support by scheduling 1-on-1 meetings with their teachers, while continuing their Universal Remote Learning schedule as planned. Meals or snacks were provided during this phase and transportation was provided for Pre-K through second grade students attending a once-a-week in-person session[3]. Based on the district’s school reopening plan, the district skipped Phase II and implemented Phase III weeks after the first phase began. Phase III and the rest of the phases were implemented for all students through the rest of the fall semester ending December 18, 2020.

Before Phase III began on September 21, 2020, a first round of parental survey was conducted to gather information on parents’ preference on children’s learning mode (in-person or remote) and mode of transportation to/from school for the rest of the phases. Parents and/or guardians were required to select option for each child between September 14 through 18, and if a parent/guardian had not made a selection for each of their kids by the end of the survey period, the default selection would be face-to-face. While it was encouraged to make a semester commitment through December 18, parents and guardians were available to retake the survey as long as the final decision is

[2]Details on the districts’ school reopening plan matrix is tabulated in Appendix A1.

[3]Transportation was provided for students in grades 3-5 and all middle/high school students returning for face-to-face instruction during future phases.

made by September 18. Students in grades 2 through 12 received a device issued by the school district, meals were provided at no charge for all students. Schools in the district had a thorough plan to meet the parents' desire to stay remote or come in face-to-face, so there seem to be little to no institutional constraint regarding provision of desired learning mode. In Phase III and Phase IV (which started on October 5), a full day and 2 days in-person instructions, respectively, were provided for all students until schools re-opened school buildings for students to voluntarily attend either full-time face-to-face or full-time Universal Remote Learning. Although parents and students self-selected into different learning modes, two factors contributed some exogenous variation in student exposure to remote learning. First, testing windows for formative assessments are fairly broad, so the dates at which individual students take exams varied widely as tabulated in Table 2. Given the phase-in of in-person learning, this translates into different exposure to remote learning between fall and winter assessments in SY 2020−21. Second, once full-time in-person instruction resumed on October 14, any student who was sick, had a fever, tested positive for COVID-19, or had been exposed to COVID-19 was expected to stay home and follow public health protocols before returning to school. Thus, differences in exposure to COVID-19 generated additional variation in the proportion of time spent in remote learning. This exogenous variation in exposure to remote learning provides an opportunity to compare outcomes of students that underwent distinct changes in instructional mode as well as to investigate whether gender differences in achievement growth varied by learning mode.

2.2 Gender Achievement Gap in Metro-Atlanta School District

There has been considerable prior research that shows girls on average outperform boys on reading/ELA exams whereas they either perform similarly or girls slightly outperform boys on math exams (Duckworth and Seligman, 2006; Lai, 2010; Sartain et al., 2023; for a meta-analysis: Voyer and Voyer, 2014). In this subsection, I briefly document the pre-existing gender achievement gap in the metro-Atlanta school district I study in this paper. In the district, a formative, adaptive assessment called iReady Diagnostic (produced by Curriculum Associates) is administered every fall and winter of each academic year; I plot standardized iReady math and reading assessment score trends to investigate pre-pandemic gender-based achievement differences in the district[4]. Figures 2a and 2b show the math and reading iReady achievement trends of the district over the analyses period (SY 2018-19 to SY 2020-21). The left side of the vertical dashed line in each figure shows the math and reading achievement trends during the pre-pandemic semesters, and the during-the-pandemic semesters trends are the right side of the dashed line in each figure. Both math and reading achievement trends for boys and girls are in line with the descriptive evidence from the gender gap literature, that girls outperform boys in both subjects where the achievement gap is wider in reading. Also, the gender-based gaps in achievement of both math and reading widen between fall and winter of SY 2020-2021, which is the period students returned to in-person learning. Although, it is important to note that these graphs do not indicate any causal relationship, but rather display the raw scale score trends of boys and girls over the analyses period.

3 Methodology

3.1 Conceptual Framework

Based on the traditional education production function, there are various inputs that potentially affect student outcomes (student academic achievement, most commonly) such as student input, family input, peer input, school and teacher input, where the function provides direct evidence about the effectiveness of each input and numer-

[4] For more details on the iReady Diagnostic, visit: https://www.curriculumassociates.com/programs/i-ready-assessment/diagnostic

ous policies that were implemented based on its estimation (Hanushek, 2020). Following the mathematical presentation of Boardman and Murnane (1979), Hanushek (1979), Todd and Wolpin (2003), and Sass et al. (2014), such relationship can be expressed as a simplified cumulative achievement function:

$$A_{it} = f(S_i(t), P_i(t), X_i(t), F_i(t), I_{i0}, \epsilon_{it}) \tag{1.1}$$

where A_{it} is a student i's academic achievement at time t, $S_i(t)$ is school-related inputs such as the number of students per school, school characteristics, teacher's experience, teacher's salary, cumulative to time t. Likewise, $P_i(t)$ is cumulative peer inputs, such as peers' academic achievement, income and socioeconomic status of peers' parents, peers' disruptiveness, $X_i(t)$ is cumulative individual/student inputs such as innate skill endowments, cognitive and non-cognitive skills such as critical thinking, consciousness, and self-discipline, and $F_i(t)$ is cumulative family-related inputs such as parents' occupation, parent's education, household income, the number of siblings, and so on. I_{i0} and ϵ_{it} are the student i's endowed innate ability and an idiosyncratic error term at time t. Taking this cumulative achievement function and the history of all inputs in time t and $t-1$ and rearranging them under several model assumptions produce the following cumulative achievement equation:

$$A_{igst} = \beta_1 X_{igst} + \beta_2 P_{-igst} + \beta_3 S_{igst} + \theta A_{igst-1} + \rho_i + \lambda_g + \sigma_s + \xi_{igst} \tag{1.2}$$

where A_{igst} is an academic achievement of a student i of grade g, in school s in year-semester t, P_{-igst} is characteristics of the student i's peers, and S_{igst} is time-varying school and teacher inputs. A_{igst-1} is a prior academic achievement of the student i, which is assumed to serve as a sufficient statistic for all prior school inputs. ρ_i, λ_g, and σ_s are time-invariant student/family, grade, and school/teacher inputs, respectively. As schools switched their learning mode from traditional face-to-face instruction to remote instruction after the pandemic broke out, the pandemic-induced school closures and the consequent shift in learning mode are believed to affect a range of educational inputs that are relevant for the process of skill formation of children (Werner and Woessmann, 2022). As aforesaid, the pandemic-induced remote learning likely changed the relative importance of educational inputs: student & family inputs, peer inputs, school & teacher inputs. Compared with the traditional face-to-face learning environment, students are less exposed to their peers and teachers as students and teachers are away from physical school buildings and classrooms. Effective self-regulated learning and parental support and supervision now become important factors to succeed in remote learning after the initial school closure, which increase the relative importance of student inputs X_{igst} (such as self-control and self-discipline skills among others) and family/household inputs ρ_i (time and resources spent on kids during remote learning, for instance), whereas such transition to remote learning would decrease the relative importance of peer inputs P_{-igst} and school/teacher inputs S_{igst}[5]. These pandemic-engendered shifts, hypothetically, would result in a relative increase in the absolute value of the coefficient on student inputs (β_1) and a relative decrease in the absolute value of the coefficient on peer inputs and school/teacher inputs (β_2 and β_3).

3.2 Data

I combine multiple administrative datasets from a metro-Atlanta school district of the period between SY 2018-2019 and SY 2020-2021, provided by the Metro Atlanta Policy Lab for Education (MAPLE) and its school district partners. The student-level panel dataset consists of rich information on student characteristics such as demographics, free or reduced-price meals (FRPM) status, English Learner (EL) status, types of disability, and mathematics and reading formative assessment scores, which are outcome

[5]In the traditional cumulative achievement function, it is assumed that family inputs are time-invariant.

variables. Two control variables of interest – proportion of historically disruptive peers and students' own self-control level – are constructed by linking the main panel dataset with Student Class and Student Discipline data; details on construction of the key variables are provided below. I restrict my sample to students in grades 1 through 8 who attended public schools in the district during the analyses period.

The first key variable of interest – proportion of historically disruptive peers in classroom – is constructed by linking the Student Class and Student Discipline data. The Student Class file includes information on which classes students took in each semester of the analysis period, and Student Discipline is student-incident-level data containing information on the type and intensity of each disciplinary incident. I link the Class and Discipline datasets to identify disruptive peers in each math and reading classes students were enrolled in and track the history of disruptiveness of their peers in each classroom. A student is considered historically disruptive if the student committed disciplinary incidents any time prior to the onset of the pandemic (but fall of SY 2018-2019 onwards), and if the type of incident falls into one of the following disciplinary codes: bullying, fighting, sexual battery, sexual harassment, sex offenses, threat or intimidation, carrying weapons (knife, handgun, rifle) and other firearms, serious bodily injury, disorderly conduct, student incivility[6]. To construct the peer disruptiveness variable, I first calculate the proportion of historically disruptive peers of all math and reading classes that each student enrolled:

$$prop.d_{icgst} = \frac{\sum_{p \neq i} P_{pcgst}}{n_{cgst} - 1} \tag{2}$$

P_{pcgst} is an indicator which equals 1 if a student i's peer p in classroom c is identified as historically disruptive in pre-pandemic period, and n_{cgst} is the number of students in the classroom c. Dividing the total number of disruptive peers ($\sum_{p \neq i} P_{pcgst}$) by the class

[6]For detailed information on disciplinary codes and frequency of each disciplinary incidents by student in the study sample, refer to Appendix C1 and C2.

size $n_{cgst} - 1$ (excluding the student i) gives us the proportion of historically disruptive peers in each classroom c, and then I calculate the average of $prop.d_{icgst}$ for math and reading courses separately to obtain the average proportion of disruptive peers in math and reading classes for the student i ($prop.d_{igst}$).

The second key variable of interest – student's self-control level – is proxied by using a "rush flag" in the main panel data[7]. Students are flagged as "being a rusher" on each of the math and reading formative assessments, where being a rusher means that a student's average time on each task of the exam were shorter than a designated time[8]. I construct the self-control variable as a dummy variable which equals 1 if students ever rushed in the exams any time in the pre-pandemic semesters.

Given that parents had options to choose between sending their kids back to school and staying remote in fall of SY 2020-2021, I employ additional data in order to conduct the analyses for the planned remote learning period, which is between fall and winter exams in SY 2020-2021. First, Blended Learning attendance data during SY 2020-2021 provide information on how many instructional days students attended under each learning mode[9]. To identify the number of attended days which students attended either remotely or face-to-face, I use the district's Blended Learning attendance data for the fall of SY 2020-2021. I calculate the proportion of remotely attended instructional

[7]Zamarro et al. (2020) take a similar approach, using item non-response and careless answering on surveys to serve as a proxy for grit and self-control. Among a sample of high school students, they find that both item non-response and careless answering were negatively correlated with both self-reported and teacher-reported measures of grit and self-control. Similarly, using data from a nationally representative panel of American adults, Zamarro et al. (2018) found that repeated careless answering behavior was negatively correlated with self-reported grit and self-reported conscientiousness. See also Hitt, Trivitt and Cheng (2016) and Hitt (2015), who study the relationship between survey effort and teacher reports of students' skills, academic outcomes at the end of high school, and college attendance.

[8]A student was given either a "yellow" flag or a "red" flag, indicating the student took less than 21 or 12 seconds on average to finish each task on the exams.

[9]Out of five partner school districts of MAPLE, only the school district I study in this paper had detailed Blended Learning data during SY 2020-2021 available. Given that detailed information on how many instructional days a student spent on each learning mode is imperative for conducting "planned blended learning" phase analysis, only students from this district are included in the analyses sample.

days as follows:

$$prop.r_{igst} = \frac{cum.remote.days_{igst=winter2021} - cum.remote.days_{igst=fall2020}}{cum.attend.days_{igst=winter2021} - cum.attend.days_{igst=fall2020}} \tag{3}$$

where $cum.remote.days_{igst}$ and $cum.attend.days_{igst}$ are cumulative attended "remote" days and "total" cumulative attended days of student i in grade g, in school s in year-semester t, respectively. $prop.r_{igst}$ is the proportion of remote learning days between the fall and winter formative assessments of SY 2020-2021, which is calculated by dividing the number of remote attendance days by total attendance days. For instance, if $prop.r_{igst}$ is 0.6 for a student i attending a school s, 60 percent of attended days between the fall and winter exams were remote[10]. Lastly, I use district's Parental Survey data – which contains information on parents' preferences toward instructional modes and types of transportation to/from school in SY 2020-2021 – and the number of COVID-19 positive and quarantined cases by school to instrument the proportion of remote learning days, to overcome selection bias issue raised by parental choice on learning mode.

Table 3 reports descriptive statistics for students in the analyses sample. Table 3a shows the statistics for full sample (columns 1-2) and subgroups by gender (columns 3-6). 42 percent of the students in the analyses sample are Black, 25 percent are White, 12 percent are Asian, and 16 percent are Hispanic. 39 percent of the students of the sample were eligible for free or reduced-price meals, FRPM, 12 percent were students with disabilities, including 8 percent of girls and 15 percent of boys. Between fall of SY 2020-2021 and winter of SY 2020-2021, students spent an average of 56 percent of attended instructional days in math remotely and 57 percent of attended instructional days in reading learning remotely. There was considerable variation in exposure to remote learning, the standard deviation in proportion remote being 0.34 in both math and reading instruction. In the "Peer Composition and Self-Control" panel in the table, I provide mean

[10]Since students take exams over a period of several weeks per each semester, the cumulative days variables are unique to each student.

statistics of variables related to students' own disruptiveness, peers' disruptiveness, and self-control. 8 percent of the students in the sample committed one or more designated disciplinary incidents during the three pre-pandemic semesters (and were thus considered "historically disruptive"). Students' own disruptiveness varied widely by gender; 4 percent of girls and 12 percent of boys were identified as disruptive students. The mean proportion of historically disruptive peers in classrooms is 8 percent for both math and reading courses. On average, 14 percent of students were an "ever-rusher" on math exams during the pre-pandemic semesters, and 10 percent were an ever-rusher on pre-pandemic reading exams. Boys were 1.8 times more likely to rush on math exams and 2.2 time more likely to rush on reading exams than girls at any point during the pre-pandemic periods. The mean statistics of our outcome variables of interest, math and reading formative assessment scores, are reported in the "Dependent Variables" panel. Girls outperform boys both on math and reading exams, where the achievement gaps are much wider in reading than in math.

As a way to check that my analyses results are not primarily driven by any changes in student/test takers composition, I break down the analysis sample by analyses period (2 semesters of pre-pandemic period, unplanned remote learning period, and planned remote learning period) and compare mean statistics of the main independent and dependent variables of each period. The statistics are reported in Table 3b. The table essentially provides the information on peer composition, self-control, and the iReady formative assessment scores of cohorts of grade 1 through 8 in each period. Compared to the pre-pandemic cohorts ("Pre-Pandemic" column), during-the-pandemic cohorts were equally or slightly more exposed to historically disruptive peers. As for the self-control level, which is proxied by rush indicator, students in the during-the-pandemic cohorts rushed more and more of them rushed at any point during the pre-pandemic semesters, comparing to those in the pre-pandemic cohorts.

3.3 Empirical Models

The analyses in the study are threefold. First, as aforementioned, I conduct an exploratory analysis to investigate the pre-pandemic relationship between being in a classroom with historically disruptive peers and academic achievement and examine whether classroom disruption was particularly problematic for girls prior to the pandemic. An analogous analysis is done with respect to student self-control level. Second, I investigate whether unplanned, emergency remote learning during the remainder of SY 2019-2020 and planned remote learning in the fall semester of SY 2020-2021 led to changes in gender-based achievement gaps. To distinguish between the effect of disruptive peers and self-control mechanisms, I allow for differential impacts based on the classroom peers' history of disruptiveness and prior measures of students' proclivity to rush. Lastly, I analyze whether gender-based academic achievement gaps differed by learning mode in Fall of SY 2020-2021.

First, I estimate the following equation over the two testing periods prior to the pandemic outbreak, Fall and Winter of SY 2019-2020 in order to explore the pre-pandemic relationship between the two mechanisms of interest and student achievement:

$$
\begin{aligned}
y_{igst} = \beta_0 + \beta_1 female_i + \beta_2 prop.d_{igst} + \beta_3 ever.rush_{igst} + \beta_4 y_{igst-1} + \beta_5 y^2_{igst-1} + \\
\beta_{\mathbf{6}} \mathbf{X_{igst}} + \lambda_g + \sigma_s + \tau_t + \epsilon_{igst} \qquad (4)
\end{aligned}
$$

where y_{igst} is standardized math and reading formative assessment scores of a student i in grade g, in school s in the beginning of pre-pandemic year-semester t[11]. $female_i$ is an indicator for female students, $prop.d_{igst}$ is the proportion of historically disruptive peers in classroom c that student i belonged during a semester before t, $ever.rush_{igst}$ is an indicator that identifies students that ever rushed in any previous pre-pandemic semesters

[11]Both math and reading formative assessment scores are standardized by grade and year-semester within the district since I could not obtain national means and standard deviations of the exams. Therefore, the analyses results should be interpreted as a within-district context.

during the sample periods[12]. In other words, $ever.rush_{igst}$ measures the student's history of proclivity to rush during the exams; it is 1 if student i ever rushed in formative assessments in any time during all previous semesters in the analyses period or 0 otherwise. y_{igst} is a prior assessment score, and $\beta_6 \mathbf{X_{igst}}$ is a vector of time-varying individual characteristics, such as the FRPM eligibility, EL status, disability status, and other relevant factors. Finally, λ_g, σ_s and τ_t refer to grade, school, and year fixed effects respectively. I run the same model without the female indicator ($female_i$) and estimate the coefficients separately for boys and girls to examine whether and to what extent classroom disruption and self-control were particularly problematic for female students prior to the pandemic.

To investigate the role of unplanned and planned remote learning in changing gender achievement gaps through the two potential mechanisms, I employ the Kitagawa-Blinder-Oaxaca decomposition method, which was first introduced in the economics literature by Ronald Oaxaca and Alan Blinder to assess the sources of male-female wage differentials (Oaxaca, 1973; Blinder, 1973; Jann, 2008). Here I use the decomposition method to decompose differences in formative assessment scores between girls and boys into three parts that could potentially explain the mean differences: (i) group differences in characteristics (e.g. the level of self-control), (ii) group differences in the marginal effects of characteristics on test scores (e.g. the impact of an increase in the proportion of historically disruptive peers on math scale scores), and (iii) differences in unobserved factors (e.g. average differences in motivation or interest in school). Generally, the third factor is attributed to discrimination in the literature on racial wage differentials. The difference in average formative assessment scores between girls (G) and boys (B) due to constant unmeasured factors, peer influences, and lack of self-control,

[12] As for $prop.d_{igst}$, I look at the history of disruptiveness of peers in a classroom in year-semester $t-1$ because the standardized tests are administered in the beginning of each semester. For example, if a dependent variable is the standardized test score in the beginning of fall of SY 2019-2020, $prop.d_{igst}$ is calculated using the history of disruptiveness of winter of SY 2018-2019 classroom peers.

$\triangle\overline{y}_t$, can be decomposed as follows:

$$\begin{aligned}\triangle\overline{y}_t &= \overline{y_t^G} - \overline{y_t^B} \\ &= (\beta_0^G - \beta_0^B) + [\alpha_1(\triangle\overline{prop.d_t}) + \alpha_2(\overline{prop.d_t^B}) + \alpha_2(\triangle\overline{prop.d_t})] + \\ &[\alpha_3(\triangle\overline{ever.rush_t}) + \alpha_4(\overline{ever.rush_t^B}) + \alpha_4(\triangle\overline{ever.rush_t})]\end{aligned} \tag{5}$$

$\triangle\overline{prop.d_t}$ is the difference between girls and boys in mean proportions of disruptive classroom peers, i.e. $\overline{prop.d_t^G} - \overline{prop.d_t^B}$, and $\triangle\overline{ever.rush_t}$ is the difference between girls and boys in the mean proportions of past rushing. $(\beta_0^G - \beta_0^B)$ represents the time-constant difference in outcomes for girls and boys that is due to unobserved gender differences not measured by explanatory variables in equation (4). α_1 is the marginal effect of disruptive peers on boys, β^B, and α_2 is the difference in the marginal effect of disruptive peers on girls and boys, $\beta^G - \beta^B$. α_3 and α_4 represent the same components of the rushing behavior indicator. The first bracketed term is the difference in outcomes between girls and boys that is due to the influences of disruptive peers. It has three components: differences due to differences in exposure to disruptive peers $(\alpha_1(\triangle\overline{prop.d_t}))$, differences due to differences in the marginal effect of disruptive peers $(\alpha_2(\overline{prop.d_t^B}))$, and the interaction of differences in the marginal effect and differences in exposure $(\alpha_2(\triangle\overline{prop.d_t}))$. The second bracketed term represents the difference in outcomes due to lack of self-control, with components analogous to those for peer influences.

In order to understand how unplanned remote instruction affected gender differences in outcomes, I estimate the decomposition model using equation (4) focusing on the period from the winter of SY 2019-2020 (just prior to the pandemic) to Fall of SY 2020-2021, after 9-weeks of unplanned remote learning in the remainder of SY 2019-2020 but before any significant return to in-person learning in SY 2020-2021. I concentrate on how the bracketed terms in equation (5) change, relative to the pre-pandemic period. Given the switch to remote learning was unplanned and there was no parental choice over learning mode, the class peer composition should not have changed significantly from prior periods. Likewise, the past proclivity of boys and girls to rush through exams should not have changed. Further, the interaction component in the first bracketed term, which is the product of two changes, should be small. Consequently, the key items of interest are changes to the marginal effects of disruptive peers on boys and girls (α_1 and $\alpha_1 + \alpha_2$), and changes to the marginal effects of lack of self-control on boys and girls (α_3 and $\alpha_3 + \alpha_4$). If remote learning dampens peer influences, we would expect either the absolute value of α_1 to decrease or the effect fade away, though the change in the difference in marginal effects between girls and boys (α_2), is unclear, a priori. If remote learning requires greater self-control, then the absolute value of the marginal effects of prior "rushing" (α_3) should increase. Even if the difference in marginal effects (α_4) does not change, the gender difference in outcomes would change if α_3 changes from the pre-pandemic period to the unplanned-remote-learning period (the term $\alpha_3(\triangle\overline{ever.rush_t})$ in equation (5) would increase). Thus, if girls have greater self-control on average than to do boys ($\triangle\overline{ever.rush_t} < 0$), then the unplanned shift to remote learning would increase gender achievement gaps assuming prior rushing has a negative effect on test scores for both boys and girls.

As discussed above, several metro-Atlanta school districts began to offer in-person instruction at varying times during SY 2020-2021, while maintaining remote learning as an option. Given that parents could choose the learning mode option for their child, this likely lead to changes in the peer composition of both in-person and remote classrooms. The marginal effects of peer composition and own self-control would also vary with learning model choice. To measure these changes and their corresponding impact of gender achievement growth differentials, I re-estimate the decomposition model for the planned-remote-learning period (between the fall and winter exams in SY 2020-2021), allowing the coefficients on each of the terms including $prop.d_{igst}$, and $ever.rush_{igst}$ to vary by the proportion of time between exams spent in remote learning. Although there are several factors that contributed to exogenous variation in exposure to remote learning, as described in subsection 2.2, the fact that learning mode is a function of parental preferences and beliefs about what learning environment is best for their child potentially complicates matters, as unmeasured factors determining learning mode may also impact student outcomes. To further address this selection bias, I employ the 2SLS method when estimating the decomposition model, using parents' preferences toward face-to-face learning and information on COVID-19 quarantined cases as instruments to predict the proportion of days spent in remote learning. This allows us to causally estimate the impact of the variation in the proportion of instructional days spent in remote learning on gender achievement gaps. Specifically, I estimate the following first-stage and the reduced-form equations:

$$prop.r_{igst} = \gamma_0 + \gamma_\mathbf{1}\mathbf{z_{igst}} + \gamma_\mathbf{2}\mathbf{X_{igst}} + e_{igst} \tag{6.1}$$

$$\begin{aligned} y_{igst} = \beta_0 + \beta_1 female_i + \beta_2 prop.d_{igst} + \beta_3 ever.rush_{igst} + \beta_4 \widehat{prop.r_{igst}} + \\ \beta_5 y_{igst-1} + \beta_6 y^2_{igst-1} + \beta_\mathbf{7}\mathbf{X_{igst}} + \lambda_g + \sigma_s + \tau_t + \epsilon_{igst} \end{aligned} \tag{6.2}$$

$prop.r_{igst}$ is the proportion of attended instructional days in remote of student i in grade g in school s between fall and winter exams in SY 2020-2021, $\mathbf{z_{igst}}$ and $\mathbf{X_{igst}}$ are a vector of instruments and covariates, respectively. The parameter $\gamma_\mathbf{1}$ in equation (6.1) captures the first-stage effect of $\mathbf{z_{igst}}$ on $prop.r_{igst}$, adjusting for controls, $\mathbf{X_{igst}}$. The estimated first-stage fitted value of $prop.r_{igst}$ ($\widehat{prop.r_{igst}}$) is then used in the equation (6.2) to estimate the effect of the two mechanisms of interest and the proportion of remote days driven by parents' preference on face-to-face learning and the pandemic-induced quarantined cases by school on student achievement and gender achievement gaps (with appropriate adjustments to the variance-covariance matrix, given that, $\widehat{prop.r_{igst}}$ is an estimate of $prop.r_{igst}$). y_{igst} represents standardized math and reading formative assessment scores of a student i of grade g in school s in the beginning of Winter of

SY 2020-2021, which is after students went through planned blended learning in the middle of fall of SY 2020-2021, and y_{igst-1} is the formative assessment scores for both subjects in the beginning of fall of SY 2020-2021, which is before students enrolled in planned blended learning. λ_g, σ_s, and τ_t refer to time-invariant grade, school-zone, and year-semester fixed effects respectively. Then I analyze whether gender-based academic achievement gaps differed by learning mode in between fall and winter of SY 2020-2021. I run two separate decomposition models, one for the "remote learning" sample and another for the "in-person learning" sample, where the remote learning sample consists of students whose proportion of remote learning days was greater than or equal to 50 percent, and the rest of the students comprise the in-person learning sample.

While I instrument the proportion of remote learning with the parental preference to face-to-face learning and the pandemic-induced quarantine cases in order to overcome the selection bias issue, the endogeneity of the parental preference of face-to-face leaning remains as a key threat to the validity of the instruments and the analyses results. Also, given that the analyses period of the unplanned, emergent remote learning include summer of SY 2019-2020 as well, the estimates might pick up the impacts of the nine-week unplanned remote learning as well as the following summer.

4 Results

4.1 Pre-Pandemic Relationship between Disruptive Peers, Self-Control Level, and Student Achievement

Before diving into the main analyses results, I first report estimates of the pre-pandemic relationship between the two mechanisms of interest and student achievement as well as the gender disparities in math and reading achievement scores prior to the onset of the pandemic. Table 4 shows the OLS analyses results for the full sample and by gender, where I report the pre-pandemic relationship between students' achievement (in terms of standardized mathematics and reading formative assessment scores) and 1) proportion of historically disruptive peers in classroom and 2) students' past proclivity to rush.

Table 4a reports OLS estimates of the full sample by different model identifications. Each column displays the estimated coefficients and robust standard errors (in parentheses) of each model specification. Model (1) only controls for *proportion of disruptive peers, ever rushed*, and the female indicator, Model (2) includes all other controls such as student demographics and characteristics, household's socioeconomic status measured by FRPM eligibility, and prior achievement. Model (3) includes grade, school, and year-semester fixed effects as well as all other controls that were previously included in Model (2). The estimated effects on *proportion of disruptive peers* and *ever rushed* across all the model identifications confirm prevalent beliefs that classroom disruptiveness and lack of self-control have negative impacts on student achievement on average. In Model (3), which is my preferred specification, I find that a 10 percentage point increase in the proportion of historically disruptive peers in classrooms decrease math (reading) formative assessment scores by 0.02 (0.018) standard deviations and being an "ever-rusher" decreases the math (reading) scores by 0.04 (0.08) standard deviations. Female students slightly outperform boys in reading, while underperform them in math.

I then re-estimate the pre-pandemic models of student achievement, separating the sample by gender; estimation results are reported in Table 4b. The magnitudes of the effect of being exposed to historically disruptive peers in classrooms range from -0.17 SD to -0.20 SD for both boys and girls, depending on test subject. The estimates indicate that girls and boys were equally affected by historically disruptive peers prior to the pandemic. As for the effect of past proclivity to rush, the magnitudes of the impact are slightly larger for boys on both subjects. For example, ever being a rusher in the past decreases boys' math achievement by 0.05 SD, whereas girls' math achievement decreases by 0.03 SD.

4.2 Unplanned Remote Learning and Planned Remote Learning

Next, I present results for the Kitagawa-Blinder-Oaxaca decomposition models in Table 5 and Table 6. In Tables 5a and 5b, I provide a short version of decomposition where I decompose the math and reading gender achievement gaps into the role played by each set of controls: proportion of disruptive peers, self-control, previous achievement, proportion of remote learning, et cetera. The top row ("Gender Achievement Gap") in both tables shows total math (reading) gender achievement gap, and the next three panels report the share of gender gaps due to (i) mean differences, (ii) differences in marginal effect, and (iii) interaction between the two effects. In both tables, the first column shows the estimates for the pre-pandemic semesters analyses (between winter of SY 2018-19 and winter of SY 2019-20), second and third columns show the estimation results for the unplanned remote learning (between the initial school closure and fall of SY 2020-21) and planned remote learning (between fall and winter of SY 2020-21), respectively. On math exams, gender achievement gaps between girls and boys were 0.03 SD in the pre-pandemic semester (fall and winter of SY 2019-20) where 5 percent of the gender gap can be explained by the difference in the mean proportion of historically disruptive peers between girls and boys. The difference in the mean statistics in the ever-rushed indicator can explain the gender gap about 11 percent. Note there are negative contributions of several components; for instance, the difference in marginal effect of previous math achievement scores ("Previous Achievement" in the second panel) plays a "negative" role on the total gender achievement gap and its share of total gender gap is 27 percent. In other words, removing the marginal effect differences of the previous achievement between girls and boys would widen the disparity by 27 percent. The time-constant difference in the math achievement due to unobserved gender differences ($\beta_0^G - \beta_0^B$) makes up a large share of the estimated total gender gap, and the differences in marginal effect of the proportion of remote learning days between girls and boys also explain a considerable share of the total gender achievement gap. Although, on reading results, mean differences in previous reading achievement and the time-constant difference due to unobserved factors have the greatest contribution to the gender gap, especially after the pandemic broke out.

Table 6 provides a detailed decomposition, where three components from the decomposition analyses are reported: (i) overall gender achievement gaps (which are also reported in Table 5), (ii) share of gender gaps attributable to each of the three components described in the equation (5) as well as the gaps explained by mean differences in the proportion of historically disruptive peers and in the propensity to rush in the past, and (iii) marginal effects of the proportion of disruptive peers and the past proclivity to rush for both boys and girls (β^B and β^G)[13]. The estimation results for standardized math and reading assessment scores are shown in Tables 6a and 6b. As previously reported in Table 4b, β^B and β^G reported in Table 6a indicate that girls and boys were equally affected by historically disruptive peers and boys are more negatively affected by lacking self-control comparing to girls prior to the pandemic. As students went through the emergent shift to remote learning in the remainder of SY 2019-20, the gender achievement gap nearly doubles in math. However, the mean differences in the proportion of historically disruptive peers no longer explain any share of the gender gap and the proportion explained by differences in self-control diminish from 11% to 9%. The changes in the magnitudes of the marginal effects of disruptive peers and self-control are discernible; disruptive peers do not significantly impact boys and girls during the unplanned remote learning comparing to the pre-pandemic period, though both boys and girls become more vulnerable to the lack of self-control throughout the unplanned remote learning period. In the last two columns, the results for the

[13]As a main model identification, the analysis is conducted with the sample consisting of distinct student groups over the course of the pandemic. To make sure that the variation in the impacts across time periods is not being driven by any changes in sample composition, I limit the sample to students appear in the entire sample periods as an alternative identification [and results. . .]

same components are shown for the planned remote learning period spanning from the beginning of fall and winter semester of SY 2020-21. Both the OLS and IV results report closing gender achievement gap in math, while the gap remain larger than its pre-pandemic level. The mean differences in the proportion of historically disruptive peers between girls and boys still do not explain any share of the gender gap, whereas the mean differences in the self-control level between girls and boys explain 13 to 15 percent of the gender gap, depending on the model identification. The magnitudes of the marginal effects of disruptive peers returns to its pre-pandemic level (even worse for boys) and the negative marginal effects of the lack of self-control stay at its unplanned remote learning level. Table 6b provides the corresponding components and results of the decomposition analyses for reading assessments. Results for reading exams follow similar patterns with those of math exams, though magnitudes of the share of the gap explained by the mean differences in self-control level are much smaller than those of math. Gender gap achievement first widens and then close over the course of the remote learning since the initial school closure, and the mean differences in proportion of disruptive peers between girls and boys become less attributable after the initial shift to the blended learning.

Lastly, tables 7a and 7b show the analogous decomposition results for math and reading, where the analyses sample during the planned remote learning in between fall and winter of SY 2020-2021 is broken down into remote and in-person groups as described earlier. Note the striking differences in math achievement gaps between students in the remote group and those in the in-person group in both OLS and IV models. The gender gaps in math achievement for those in remote instruction group widened even more over the course of the pandemic, where the gaps are insignificant for those in in-person group. The mean differences in the self-control level between girls and boys explain 6 percent of such widened gaps, where the mean differences in the proportion of disruptive peers hardly explain such gaps. For reading, in contrast, the achievement gaps between girls and boys are slightly wider for those in the in-person group. The estimated marginal effects of the proportion of historically disruptive peers on both math and reading indicate that students in the remote group are more vulnerable to disruptive peers than those in the in-person group.

5 Discussion & Conclusion

The COVID-19 pandemic and the school closures has undoubtedly affected many aspects of people's lives. Especially for students, the pandemic-induced shift to remote learning has unprecedentedly altered the nature of their learning environment: being away from school buildings, peers, and teachers, and learning from home under parents' and guardians' supervision. Since the initial onset of the COVID-19 pandemic, remote learning has received great attention and the importance of self-regulated learning has been stressed ever than before (Berger et al., 2021). In fact, remote learning is now commonly offered as an alternative to the traditional in-person learning; nearly all of the United States' largest school districts announced to continue providing a remote option as well as expanding their virtual learning offerings for the fall of SY 2022-2023, and several states intermittently switched to remote learning or dismiss students early in the day when cities within the states experience hot days at schools (Belsha and Barnum, 2022; Will, 2022).

In this paper, I study the impact of the pandemic-induced remote learning on student achievement and gender achievement gaps, focusing on the change in relative importance of impacts from disruptive peers and students' own self-control level due to the pandemic. The results suggest that disruptive peers and self-control continue to be significant determinants of student achievement over the course of the pandemic, and students' pre-pandemic self-control level can explain a moderate share of the observed gender achievement gaps where the gaps favor female students. While baseline score (measured by previous performance) explains majority of gender achievement gaps, gender-based differences in mean of self-control level and impact of disruptive peers and self-control level between girls and boys explain moderate share of gender achievement gaps. Moreover, I find large gender achievement gaps in math for those who stayed remote between the fall and winter exams of SY 2020-2021, whereas no statistically significant gender gaps were found among those returned to school.

The finding on the disparities in gender gaps by learning mode is in line with what Goldhaber et al. (2022) finds, that remote instruction was a primary driver of widening achievement gaps between high and low poverty schools, where math gaps did not widen in areas that remained in-person though there was some widening in reading gaps in those areas. Accumulating evidence on disproportionate impact of the pandemic-induced remote learning by gender as well as race/ethnicity and socioeconomic status raise serious concerns since such exacerbated gaps could translate into larger gaps in future outcomes, such as post-secondary school outcomes and earnings. Moreover, additional evidence on the negative consequences of remote learning versus in-person learning due to the pandemic outbreak on top of the findings from this study are also worrisome, given there has been growing supply and demand of distance learning as stated earlier (Jack et al., 2022; Tagami, 2022). Based on the findings from this paper, one way to close such exacerbated gender gaps due to the pandemic and the pandemic-induced shift in learning environment would be to devise a way to measure students' non-cognitive skills that are essential to successfully navigate through self-regulated learning, such as self-control and perseverance, and accordingly support students provided that the results indicate students' self-control level could explain up to 15 percent of the gender gaps in the district. For instance, school districts could offer several remedies to students participating or have participated in various learning modes as well as provide additional supports to students lacking self-control and self-discipline to catch upon the learning disruption caused by the pandemic. Also, given the results that the gender-based impact differences in disruptive peers in classroom and self-control contribute to the achievement gaps between girls and boys, it is suggested to identify and support students struggling in different learning modes and continue monitoring students' disciplinary behaviors given that districts are experiencing a surge in student disciplinary incidents, seemingly induced by the pandemic, along with the existing research showing clear evidence of negative impacts of disruptive peers in classrooms (Hoxby, 2000; Carrell et al., 2018; Wile, 2022; Downey, 2022; McCray, 2022)

This study is not without limitations. First, I do not observe any variables related to parents and household characteristics. While some children had affluent resources from their parents, others did not necessarily have such support at home. Moreover, it is believed that parents allocate more efforts to girls than to boys, and there is a negative correlation between parental efforts and prior achievement (Bonesrønning, 2010) Unfortunately, the analyses with respect to family inputs are beyond the scope of this paper since I cannot control for any family-related variables in the analysis reported in this paper. Nevertheless, the analyses provide the overall snapshot of what has happened over the course of the pandemic, and the results would provide valuable information to the districts, policymakers, and parents for making future decisions. While I focus on decomposing overall gender achievement gaps and exploring average impacts of the pandemic-induced remote learning of pooled sample of grades 1 to 8 students by gender, further analysis is needed to fully investigate the mechanisms of gender achievement gaps, especially to explore which student subgroups within the same gender (by race, SES, academic achievement, grade, et cetera) particularly suffered from the remote learning and consequently experienced larger achievement gaps. Ultimately, future research needs to be done to examine the long-term consequences of the pandemic-induced remote learning with respect to disruptive peer effects and self-control to learn how best to support students under various learning modes in the future.

References

Aucejo, E. M., French, J., Araya, M. P. U., & Zafar, B. (2020). The impact of COVID-19 on student experiences and expectations: Evidence from a survey. *Journal of Public Economics, 191*. https://doi.org/https://doi.org/10.1016/j.jpubeco.2020.104271

Autor, D., Figlio, D. N., Karbownik, K., Roth, J., & Wasserman, M. (2020). *Males at the tails: How socioeconomic status shapes the gender gap* (tech. rep.). National bureau of economic research.

Bacher-Hicks, A., Goodman, J., Green, J. G., & Holt, M. K. (2022). The covid-19 pandemic disrupted both school bullying and cyberbullying. *American Economic Review: Insights, 4*(3), 353–70.

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Au Yeung, N. (2021). Achievement gaps in the wake of covid-19. *Educational Researcher, 50*(5), 266–275.

Belsha, K., & Barnum, M. (2022, June 6). *Sticking around: Most big districts will offer virtual learning this fall, a sign of pandemic's effect.* Chalkbeat. https://www.chalkbeat.org/2022/6/6/23153483/big-school-districts-virtual-learning-fall-2022

Berger, F., Schreiner, C., Hagleitner, W., Jesacher-Rößler, L., Roßnagl, S., & Kraler, C. (2021). Predicting coping with self-regulated distance learning in times of covid-19: Evidence from a longitudinal study. *Frontiers in Psychology, 12*, 701255.

Bertrand, M., & Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American economic journal: applied economics, 5*(1), 32–64.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human resources*, 436–455.

Boardman, A. E., & Murnane, R. J. (1979). Using panel data to improve estimates of the determinants of educational achievement. *Sociology of education*, 113–121.

Bonesrønning, H. (2010). Are parental effort allocations biased by gender? *Education Economics, 18*(3), 253–268.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review, 108*(11), 3377–3415.

Chen, D. L., Ertac, S., Evgeniou, T., Miao, X., Nadaf, A., & Yilmaz, E. (2022). Grit and academic resilience during the covid-19 pandemic.

Copeland, W. E., McGinnis, E., Bai, Y., Adams, Z., Nardone, H., Devadanam, V., Rettew, J., & Hudziak, J. J. (2021). Impact of covid-19 pandemic on college student mental health and wellness. *Journal of the American Academy of Child & Adolescent Psychiatry, 60*(1), 134–141.

Dahl, G., & Lochner, L. (2005). The impact of family income on child achievement.

Donnelly, R., & Patrinos, H. A. (2021). Learning loss during covid-19: An early systematic review. *Prospects*, 1–9.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). *Covid-19 and learning loss - disparities grow and students need help*. McKinsey & Company. https://wasa-oly.org/WASA/images/WASA/5.0%20Professional%20Development/4.2%20Conference%20Resources/Winter/2021/covid-19-and-learning-loss-disparities-grow-and-students-need-help-v3.pdf

Downey, M. (2022, September 12). *Opinion: Pandemic plays role in rising middle school bullying, drama*. The Atlanta Journal-Constitution. https://www.ajc.com/education/get-schooled-blog/opinion-pandemic-plays-role-in-rising-middle-school-bullying-drama/Y7ICX2JQQBHRHICEUFX5IXYGAM/

Duckworth, A. L., Shulman, E. P., Mastronarde, A. J., Patrick, S. D., Zhang, J., & Druckman, J. (2015). Will not want: Self-control rather than motivation explains the female advantage in report card grades. *Learning and individual differences, 39*, 13–23.

Duckworth, A. L., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of educational psychology*, *98*(1), 198.

Ficano, C. C. (2012). Peer effects in college academic outcomes–gender matters! *Economics of Education Review*, *31*(6), 1102–1115.

Georgia Department of Education. (2020, August 1). *Covid-19 (coronavirus) and schools*. Georgia Insights of GaDOE. https://www.georgiainsights.com/georgia-school-closures-370622.html

Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2022). *The consequences of remote and hybrid instruction during the pandemic* (tech. rep.). National Bureau of Economic Research.

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of covid-19-related school closures on student achievement-a systematic review. *Frontiers in Psychology*, 4020.

Han, L., & Li, T. (2009). The gender difference of peer influence in higher education. *Economics of Education review*, *28*(1), 129–134.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of human Resources*, 351–388.

Hanushek, E. A. (2020). Education production functions. In *The economics of education* (pp. 161–170). Elsevier.

Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. *NBER Working Paper*, (w7867).

Jack, R., Halloran, C., Okun, J., & Oster, E. (2022). Pandemic schooling mode and student test scores: Evidence from us school districts. *American Economic Review: Insights*.

Jacob, B. A. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education review*, *21*(6), 589–598.

Jann, B. (2008). The blinder–oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453–479.

Krein, S. F., & Beller, A. H. (1988). Educational attainment of children from single-parent families: Differences by exposure, gender, and race. *Demography*, *25*(2), 221–234.

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test score patterns across three covid-19-impacted school years. *EdWorkingPaper: 22-521*, 37–62.

Lai, F. (2010). Are boys left behind? the evolution of the gender achievement gap in beijing's middle schools. *Economics of Education Review*, *29*(3), 383–399.

Lane, M. A. (2020, March 16). *Gov. kemp orders all k-12 georgia schools to close until end of march*. Georgia Public Broadcasting. https://www.gpb.org/blogs/education-matters/2020/03/16/gov-kemp-orders-all-k-12-georgia-schools-close-until-end-of

McCray, V. (2022, March 31). *Fulton county schools faces surge in student disciplinary issues*. The Atlanta Journal-Constitution. https://www.ajc.com/education/fulton-county-schools-faces-surge-in-student-disciplinary-issues/UKGFYYW6MVFWFIVABVEEEFZ

Nakajima, N., Jung, H., Pradhan, M., Hasan, A., Kinnell, A., & Brinkman, S. (2020). Gender gaps in cognitive and social-emotional skills in early primary grades: Evidence from rural indonesia. *Developmental science*, *23*(5), e12931.

National Center for Education Statistics. (2022, September 1). *Naep long-term trend assessment results: Reading and mathematics*. https://www.nationsreportcard.gov/highlights/ltt/2022/

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, 693–709.

OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. PISA, OECD Publishing, Paris. https://doi.org/https://doi.org/10.1787/b5fd1b8f-en

Sartain, L., Freire, S., Easton, J. Q., & Diaz, B. (2023). When girls outperform boys: The gender gap in high school math grades. *EdWorkingPaper: 23-707*. https://doi.org/https://doi.org/10.26300/b0py-tz14

Sass, T., & Goldring, T. (2021). Student achievement growth during the covid-19 pandemic. https://gpl.gsu.edu/publications/student-achievement-growth-during-the-covid-19-pandemic/

Sass, T., Semykina, A., & Harris, D. N. (2014). Value-added models and the measurement of teacher productivity. *Economics of Education Review*, *38*, 9–23.

Skar, G. B. U., Graham, S., & Huebner, A. (2021). Learning loss during the covid-19 pandemic and the impact of emergency remote instruction on first grade students' writing: A natural experiment. *Journal of Educational Psychology*.

Tagami, T. (2022, September 14). *Research suggests in-person classes helped georgia kids during covid*. The Atlanta Journal-Constitution. https://www.ajc.com/education/research-suggests-in-person-classes-helped-georgia-kids-during-covid/H3CPHVCONNB2PO62II7EZSOSAA/

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), F3–F33.

UNESCO Institute for Statistics. (2021). *Dashboards on the Global Monitoring of School Closures Caused by the COVID-19 Pandemic*. http://covid19.uis.unesco.org/global-monitoring-school-closures-covid19/country-dashboard/

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological bulletin*, *140*(4), 1174.

Werner, K., & Woessmann, L. (2022). The legacy of covid-19 in education. *EdWorkingPaper: 21-478*.

Wile, O. (2022, August 27). *Georgia department of education reports school fights were up in 2021-2022 school year*. WJCL. https://www.wjcl.com/article/georgia-department-of-education-reports-school-fights-were-up-in-2021-2022-school-year/41006554

Will, M. (2022, September 26). *The school year is getting hotter. how does heat affect student learning and well-being?* EducationWeek. https://www.edweek.org/leadership/the-school-year-is-getting-hotter-how-does-heat-affect-student-learning-and-well-being/2022/09

Zamarro, G., Nichols, M., Duckworth, A. L., & D'Mello, S. K. (2020). Validation of survey effort measures of grit and self-control in a sample of high school students. *PloS one*, *15*(7), e0235396.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics*, *85*(1), 9–23.

Tables and Figures

Table 1: Phase and Actual Timing of Return to Full-Time In-Person Instruction in the District

Phases	Learning Mode	Actual Start Date
Universal Remote Learning	All remote	First day of school
Phase I	90-minute session, once a week (Pre-K–2), 1:1 meeting by appointment (3–12)	September 9, 2020
Phase II	1 half-day f2f per week	N/A
Phase III	1 full-day f2f per week	September 21, 2020
Phase IV	2 full days f2f per week	October 5, 2020
Phase V	Full-time f2f or remote	October 14, 2020

Table 2: Descriptive Statistics of Testing Window for Fall and Winter Exams, SY 2020−21

(a) Testing Window

		Testing Window	Mean	Median
Math	Fall 2020−21	8/24/2020−10/23/2020	9/2/2020	9/1/2020
	Winter 2020−21	11/30/2020−1/31/2021	12/29/2020	1/7/2021
Reading	Fall 2020−21	8/24/2020−10/23/2020	9/1/2020	8/31/2020
	Winter 2020−21	11/30/2020−1/30/2021	12/30/2020	1/7/2021

(b) Number of Attended Days between Fall and Winter Exams, SY 2020−21

	Mean	SD	Min.	Max.
Math	67.29	9.00	22	87
Reading	68.37	9.38	24	87

Table 3: Descriptive Statistics

(a) Full Sample, by Gender

Variables	Full Sample Mean	Full Sample SD	Girls Mean	Girls SD	Boys Mean	Boys SD	Mean Difference (G-B)
Demographics							
Black	0.42	0.49	0.42	0.49	0.42	0.49	
White	0.25	0.44	0.25	0.43	0.26	0.44	
Asian	0.12	0.33	0.12	0.33	0.12	0.33	
Hispanic	0.16	0.37	0.16	0.37	0.17	0.37	
FRPM	0.39	0.49	0.39	0.49	0.40	0.49	
Any Disability	0.12	0.32	0.08	0.27	0.15	0.36	
Learning Mode between Fall and Winter of SY 2020-21							
Remote Days Proportion (Math)	0.56	0.34	0.57	0.35	0.56	0.34	0.01***
Remote Days Proportion (Reading)	0.57	0.34	0.58	0.34	0.56	0.34	0.02***
Peer Composition and Self-Control							
Any Disruptive Behaviors	0.08	0.27	0.04	0.20	0.12	0.32	-0.08***
Prop. of Disruptive Peers (Math)	0.08	0.13	0.08	0.12	0.09	0.13	-0.01***
Prop. of Disruptive Peers (Reading)	0.08	0.12	0.08	0.12	0.09	0.13	-0.01***
Rushed (Math)	0.03	0.18	0.02	0.15	0.04	0.20	-0.02***
Rushed (Reading)	0.06	0.25	0.05	0.21	0.08	0.27	-0.03***
Ever Rushed (Math)	0.14	0.35	0.10	0.30	0.18	0.38	-0.08***
Ever Rushed (Reading)	0.10	0.30	0.06	0.25	0.13	0.33	-0.07***
Dependent Variables							
iReady Scale Score (Math)	456.58	53.03	457.67	52.21	455.53	53.80	2.14***
iReady Scale Score (Reading)	538.94	82.37	545.15	80.78	532.93	83.43	12.21***
N	214,242		105,282		108,960		
(Number of Students)	(69,737)		(34,274)		(35,473)		

Notes: Sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the analyses period of SY 2019-2020 SY 2020-2021. Remote Days Proportion of Math and Reading report mean statistics of attended remote learning days between fall and winter formative assessments of SY 2020-2021. Detailed information on how students were identified as disruptive students can be found on Appendix A2. Details on Proportion of Disruptive Peers and Ever Rushed variables construction can be found in Section 3.2. The unit of the number of observations is individual in each school-year-semester, unique number of students are also reported in the last row.

(b) Full Sample, by Year-Semester

	Pre-Pandemic		Unplanned		Planned	
Variables	Mean	SD	Mean	SD	Mean	SD
Learning Mode between Fall and Winter of SY 2020-21						
Remote Days Proportion (Math)					0.56	0.34
Remote Days Proportion (Reading)					0.57	0.34
Peer Composition and Self-Control						
Any Disruptive Behaviors	0.08	0.27	0.09	0.29	0.08	0.26
Prop. of Disruptive Peers (Math)	0.08	0.12	0.09	0.13	0.08	0.13
Prop. of Disruptive Peers (Reading)	0.08	0.12	0.10	0.13	0.07	0.13
Rushed (Math)	0.03	0.17	0.04	0.19	0.04	0.20
Rushed (Reading)	0.06	0.23	0.07	0.25	0.08	0.27
Ever Rushed (Math)	0.13	0.34	0.15	0.36	0.15	0.36
Ever Rushed (Reading)	0.08	0.28	0.11	0.32	0.11	0.31
Dependent Variables						
iReady Scale Score (Math)	453.85	52.02	456.22	52.83	462.73	54.81
iReady Scale Score (Reading)	535.55	82.62	539.06	81.57	546.08	82.16
N	105,904		54,421		53,917	

Notes: Sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the analyses period of SY 2019-2020-SY 2020-2021. Remote Days Proportion of Math and Reading report mean statistics of attended remote learning days between fall and winter formative assessments of SY 2020-2021. Detailed information on how students were identified as disruptive students can be found on Appendix C1. Details on Proportion of Disruptive Peers and Ever Rushed variables construction can be found in Section 3.2. The unit of the number of observations is individual in each school-year-semester.

Table 4: Pre-Pandemic Relationship (OLS Results) by Subject

(a) Full Sample

	(1)		(2)		(3)	
	Math	Reading	Math	Reading	Math	Reading
Proportion of Disruptive Peers	-2.55***	-2.26***	-0.13***	-0.10***	-0.20***	-0.18***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Ever Rushed	-0.70***	-0.75***	-0.06***	-0.09***	-0.04***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female	-0.04***	0.13***	-0.02***	0.01***	-0.03***	0.01***
	(0.01)	(0.01)	(0.004)	(0.004)	(0.004)	(0.004)
Black (ref. White)			-0.11***	-0.11***	-0.08***	-0.07***
			(0.01)	(0.01)	(0.01)	(0.01)
Asian			0.10***	0.04***	0.08***	0.03***
			(0.01)	(0.01)	(0.01)	(0.01)
Hispanic			-0.06***	-0.08***	-0.04***	-0.06***
			(0.01)	(0.01)	(0.01)	(0.01)
FRPM			-0.05***	-0.05***	-0.05***	-0.05***
			(0.005)	(0.01)	(0.005)	(0.005)
Controls			Y		Y	
Grade FE					Y	
School FE					Y	
Year-Semester FE					Y	
N	83,102	84,011	69,763	66,943	69,763	66,943

Notes: Analyses Sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district during the pre-pandemic semesters (winter of SY 2018-2019 and fall of SY 2019-2020). Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester, so if a student was observed during the pre-pandemic semesters, there would be two observations for each student. Outcome variables are standardized math and reading achievement scores.

(b) Full Sample, by Gender

	Math		Reading	
	Boys	Girls	Boys	Girls
Proportion of Disruptive Peers	-0.20***	-0.20***	-0.17***	-0.17***
	(0.03)	(0.03)	(0.03)	(0.03)
Ever Rushed	-0.05***	-0.03***	-0.09***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Black (ref. White)	-0.10***	-0.06***	-0.09***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
Asian	0.08***	0.08***	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Hispanic	-0.04***	-0.05***	-0.05***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
FRPM	-0.05***	-0.05***	-0.05***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
Controls	Y	Y	Y	Y
Grade FE	Y	Y	Y	Y
School FE	Y	Y	Y	Y
Year-Semester FE	Y	Y	Y	Y
N	35,293	34,470	34,070	32,873

Notes: Analyses Sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district during the pre-pandemic semesters (fall and winter exams of SY 2019-2020). Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester, so if a student was observed during the pre-pandemic semesters, there would be two observations for each student. Outcome variables are standardized math and reading achievement scores.

Table 5: Kitagawa-Blinder-Oaxaca Decomposition Results, Full Sample

(a) Math

	(1)	(2)	(3)	
	Pre-Pandemic	Unplanned	Planned Remote	
		Remote	OLS	IV
Gender Achievement Gap	0.0340***	0.0654***	0.0498***	0.0414***
(G-B)	(0.0074)	(0.0109)	(0.0123)	(0.0127)
Gaps due to Mean Differences of:				
Proportion of Disruptive Peers	0.0016	0.0007	0.0006	0.0004
%	*5%*	*1%*	*1%*	*1%*
Self-Control Level	0.0038	0.0060	0.0066	0.0061
%	*11%*	*9%*	*13%*	*15%*
Previous Achievement	0.0289	-0.0065	0.0192	0.0115
%	*85%*	*-10%*	*39%*	*28%*
Gaps due to Marginal Effect Differences of:				
Proportion of Disruptive Peers	0.0004	-0.0039	0.0030	0.0029
%	*1%*	*-6%*	*6%*	*7%*
Self-Control Level	0.0021	-0.0084	-0.0041	-0.0051
%	*6%*	*-13%*	*-8%*	*-12%*
Previous Achievement	-0.0093	-0.0171	-0.0180	-0.0102
%	*-27%*	*-26%*	*-36%*	*-25%*
Proportion of Remote Learning			0.0307	0.0255
%			*62%*	*62%*
Gaps due to Interaction of Mean Differences and Marginal Effect Differences				
All Controls	0.0036	0.0073	0.0105	0.0094
%	*11%*	*11%*	*21%*	*23%*
N	69,763	32,409	26,179	23,857

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2018-2019 SY 2020-2021. Robust standard error in parentheses below estimated gender achievement gaps. The unit of the number of observations is individual in each school-year-semester. (1) is pre-pandemic period (2 semesters prior to the school closure), (2) is unplanned remote learning period (between the school closure and the remainder of SY 2019-2020), and (3) is planned remote learning period (between Fall and Winter of SY 2020-2021). Proportion of remote learning is instrumented in the IV model (column 4) to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

(b) Reading

	(1)	(2)	(3)	
	Pre-Pandemic	Unplanned	Planned Remote	
		Remote	OLS	IV
Gender Achievement Gap	0.2030***	0.2131***	0.2008***	0.1928***
(G-B)	(0.0075)	(0.0111)	(0.0114)	(0.0118)
Gaps due to Mean Differences of:				
Proportion of Disruptive Peers	0.0019	0.0011	0.0008	0.0009
%	*1%*	*1%*	*0.4%*	*0.5%*
Self-Control Level	0.0061	0.0055	0.0051	0.0053
%	*3%*	*3%*	*3%*	*3%*
Previous Achievement	0.1590	0.1291	0.1317	0.1232
%	*78%*	*61%*	*66%*	*64%*
Gaps due to Marginal Effect Differences of:				
Proportion of Disruptive Peers	0.0005	0.0016	-0.0083	-0.0078
%	*0.2%*	*1%*	*-4%*	*4%*
Self-Control Level	0.0031	0.0029	-0.0029	-0.0026
%	*2%*	*1%*	*-1%*	*-1%*
Previous Achievement	-0.0166	-0.0055	-0.0239	-0.0237
%	*-8%*	*-3%*	*-12%*	*1-2%*
Proportion of Remote Learning			-0.0083	0.0108
%			*-4%*	*6%*
Gaps due to Interaction of Mean Differences and Marginal Effect Differences				
All Controls	0.0009	0.0039	0.0126	0.0118
%	*0.4%*	*2%*	*6%*	*6%*
N	66,943	32,424	30,067	27,158

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2018-2019 SY 2020-2021. Robust standard error in parentheses below estimated gender achievement gaps. The unit of the number of observations is individual in each school-year-semester. (1) is pre-pandemic period (2 semesters prior to the school closure), (2) is unplanned remote learning period (between the school closure and the remainder of SY 2019-2020), and (3) is planned remote learning period (between Fall and Winter of SY 2020-2021). Proportion of remote learning is instrumented in the IV model (column 4) to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

Table 6: Detailed Kitagawa-Blinder-Oaxaca Decomposition Results, Full Sample

(a) Math

	(1)	(2)	(3)	
	Pre-Pandemic	Unplanned	Planned Remote	
		Remote	OLS	IV
Gender Achievement Gap	0.0340***	0.0654***	0.0498***	0.0414***
(G-B)	(0.0074)	(0.0109)	(0.0123)	(0.0127)
Proportion of Historically Disruptive Peers				
$\alpha_1(\triangle\overline{prop.d_t})$	0.0016	0.0007	0.0006	0.0004
% of Total Gender Gap	*5%*	*1%*	*1%*	*1%*
$\alpha_2(\overline{prop.d_t^B})$	0.0004	-0.0039	0.0030	0.0029
% of Total Gender Gap	*1%*	*-6%*	*6%*	*7%*
$\alpha_2(\triangle\overline{prop.d_t})$	-0.00004	0.0004	-0.0001	-0.0001
$\beta^G(\alpha_1)$	-0.20***	-0.12**	-0.20***	-0.20***
$\beta^B(\alpha_1+\alpha_2)$	-0.20***	-0.08	-0.24***	-0.24***
Pre-Pandemic Rush History				
$\alpha_3(\triangle\overline{ever.rush_t})$	0.0038	0.0060	0.0066	0.0061
% of Total Gender Gap	*11%*	*9%*	*13%*	*15%*
$\alpha_4(\overline{ever.rush_t^B})$	0.0021	-0.0084	-0.0041	-0.0051
% of Total Gender Gap	*6%*	*13%*	*-8%*	*-12%*
$\alpha_4(\triangle\overline{ever.rush_t})$	-0.0009	0.0034	0.0017	0.0021
$\beta^G(\alpha_3)$	-0.03***	-0.11***	-0.10***	-0.11***
$\beta^B(\alpha_3+\alpha_4)$	-0.05***	-0.07***	-0.08***	-0.08***
N	69,763	32,409	26,179	23,857

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2019-2020 SY 2020-2021. Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester. (1) is pre-pandemic period (2 semesters prior to the school closure), (2) is unplanned remote learning period (between the school closure and the remainder of SY 2019-2020), and (3) is planned remote learning period (between Fall and Winter of SY 2020-2021). Percent of Total Gender Gap is in square brackets if insignificant. Proportion of remote learning is instrumented to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

(b) Reading

	Reading			
	(1)	(2)	(3)	
	Pre-Pandemic	Unplanned	Planned Remote	
		Remote	OLS	IV
Gender Achievement Gap	0.2030***	0.2131***	0.2008***	0.1928***
(G-B)	(0.0075)	(0.0111)	(0.0114)	(0.0118)
Proportion of Historically Disruptive Peers				
$\alpha_1(\triangle\overline{prop.d_t})$	0.0019	0.0011	0.0008	0.0009
% of Total Gender Gap	*1%*	*1%*	*0.4%*	*0.5%*
$\alpha_2(\overline{prop.d_t^B})$	0.0005	0.0016	-0.0083	-0.0078
% of Total Gender Gap	*0.2%*	*1%*	*4%*	*-4%*
$\alpha_2(\triangle\overline{prop.d_t})$	-0.0001	-0.0002	0.0009	0.0007
$\beta^G(\alpha_1)$	-0.17***	-0.09	-0.20***	-0.22***
$\beta^B(\alpha_1+\alpha_2)$	-0.17***	-0.11*	-0.10*	-0.12***
Pre-Pandemic Rush History				
$\alpha_3(\triangle\overline{ever.rush_t})$	0.0061	0.0055	0.0051	0.0053
% of Total Gender Gap	*3%*	*3%*	*3%*	*3%*
$\alpha_4(\overline{ever.rush_t^B})$	0.0031	0.0029	-0.0029	-0.0026
% of Total Gender Gap	*2%*	*1%*	*-1%*	*-1%*
$\alpha_4(\triangle\overline{ever.rush_t})$	-0.0016	-0.0014	0.0014	0.0012
$\beta^G(\alpha_3)$	-0.07***	-0.06***	-0.10***	-0.11***
$\beta^B(\alpha_3+\alpha_4)$	-0.09***	-0.07***	-0.08***	-0.09***
N	66,943	32,424	30,067	27,158

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2019-2020 SY 2020-2021. Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester. (1) is pre-pandemic period (2 semesters prior to the school closure), (2) is unplanned remote learning period (between the school closure and the remainder of SY 2019-2020), and (3) is planned remote learning period (between Fall and Winter of SY 2020-2021). Percent of Total Gender Gap is in square brackets if insignificant. Proportion of remote learning is instrumented to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

Table 7: Kitagawa-Blinder-Oaxaca Decomposition Results by Leaning Mode, Full Sample

(a) Math

	Remote		In-Person	
	OLS	IV	OLS	IV
Gender Achievement Gap	0.0847***	0.0800***	0.0141	0.0013
(G-B)	(0.0186)	(0.0191)	(0.0162)	(0.0170)
Proportion of Historically Disruptive Peers				
$\alpha_1(\triangle\overline{prop.d_t})$	-0.0001	-0.0001	0.0013	0.0009
% of Total Gender Gap	*-0.1%*	*-0.1%*	*9%*	*69%*
$\alpha_2(\overline{prop.d_t^B})$	0.0034	0.0053	0.0030	0.0013
% of Total Gender Gap	*4%*	*7%*	*21%*	*100%*
$\alpha_2(\triangle\overline{prop.d_t})$	0.0000	0.0000	-0.0002	-0.0001
% of Total Gender Gap	*<0.1%*	*<0.1%*	*-1%*	*-8%*
$\beta^G(\alpha_1)$	-0.22***	-0.21***	-0.18***	-0.19**
$\beta^B(\alpha_1+\alpha_2)$	-0.26***	-0.28***	-0.22***	-0.21***
Pre-Pandemic Rush History				
$\alpha_3(\triangle\overline{ever.rush_t})$	0.0049	0.0044	0.0082	0.0076
% of Total Gender Gap	*6%*	*6%*	*58%*	*585%*
$\alpha_4(\overline{ever.rush_t^B})$	-0.0127	-0.0146	0.0046	0.0044
% of Total Gender Gap	*-15%*	*-18%*	*33%*	*338%*
$\alpha_4(\triangle\overline{ever.rush_t})$	0.0050	0.0059	-0.0019	-0.0018
% of Total Gender Gap	*6%*	*7%*	*-13%*	*-138%*
$\beta^G(\alpha_3)$	-0.13***	-0.14***	-0.07***	-0.07***
$\beta^B(\alpha_3+\alpha_4)$	-0.07***	-0.06***	-0.10***	-0.09***
N	12,474	11,683	13,705	12,174

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2019-2020 SY 2020-2021. Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester. Percent of Total Gender Gap is in square brackets if insignificant. Proportion of remote learning is instrumented to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

(b) Reading

	Remote		In-Person	
	OLS	IV	OLS	IV
Gender Achievement Gap	0.1917***	0.1882***	0.2046***	0.1941***
(G-B)	(0.0165)	(0.0169)	(0.0157)	(0.0164)
Proportion of Historically Disruptive Peers				
$\alpha_1(\triangle\overline{prop.d_t})$	0.0007	0.0009	0.0011	0.0011
% of Total Gender Gap	*0.4%*	*0.5%*	*1%*	*1%*
$\alpha_2(\overline{prop.d_t^B})$	-0.0153	-0.0123	0.0008	0.0006
% of Total Gender Gap	*-8%*	*-7%*	*0.4%*	*0.3%*
$\alpha_2(\triangle\overline{prop.d_t})$	0.0012	0.0010	-0.0001	-0.0001
% of Total Gender Gap	*1%*	*1%*	*<0.1%*	*-0.1%*
$\beta^G(\alpha_1)$	-0.28***	-0.27***	-0.10	-0.12
$\beta^B(\alpha_1+\alpha_2)$	-0.10	-0.13**	-0.11	-0.13*
Pre-Pandemic Rush History				
$\alpha_3(\triangle\overline{ever.rush_t})$	0.0063	0.0057	0.0042	0.0052
% of Total Gender Gap	*3%*	*3%*	*2%*	*3%*
$\alpha_4(\overline{ever.rush_t^B})$	0.0008	-0.0016	-0.0054	-0.0025
% of Total Gender Gap	*0.4%*	*-1%*	*-3%*	*-1.3%*
$\alpha_4(\triangle\overline{ever.rush_t})$	-0.0004	0.0007	0.0026	0.0012
% of Total Gender Gap	*-0.2%*	*0.4%*	*1.3%*	*1%*
$\beta^G(\alpha_3)$	-0.09***	-0.10***	-0.11***	-0.11***
$\beta^B(\alpha_3+\alpha_4)$	-0.10***	-0.09***	-0.07***	-0.09***
N	14,325	13,309	15,742	13,849

Notes: Analyses sample includes students Grade 1 to Grade 8 enrolled in public schools located in the school district over the period of SY 2019-2020 SY 2020-2021. Robust standard error in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester. Percent of Total Gender Gap is in square brackets if insignificant. Proportion of remote learning is instrumented to address selection bias in the planned remote learning period, where IVs are parent's preference for f2f and number of COVID-19 quarantines cases by school.

Figure 1: Timeline of School Closure and iReady Diagnostic Testing Windows

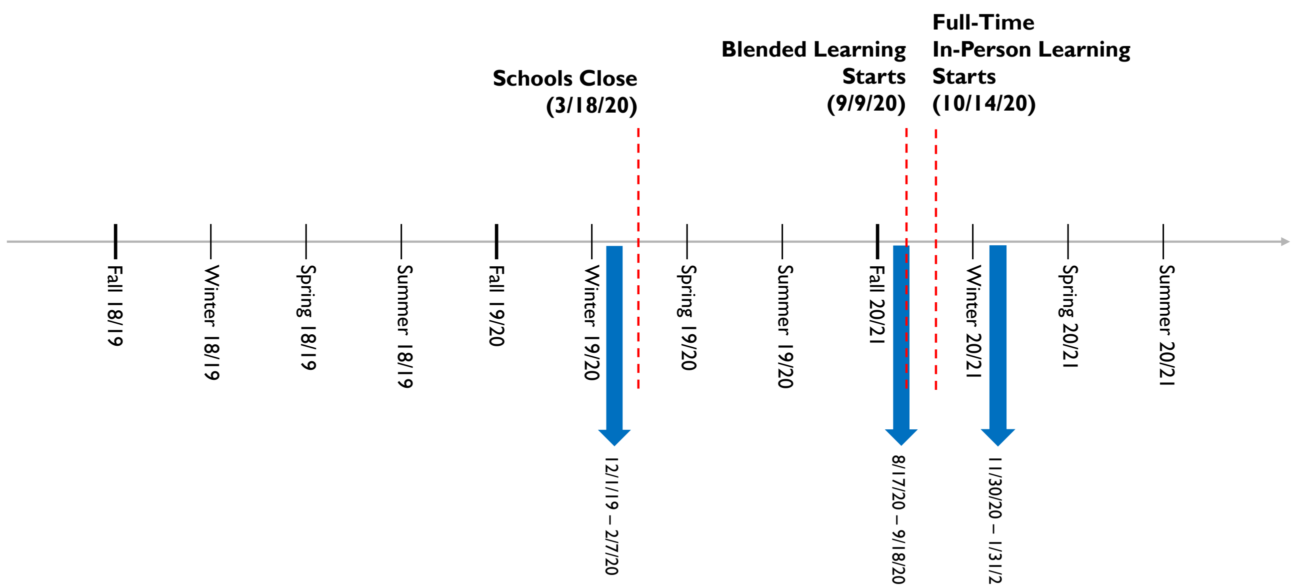
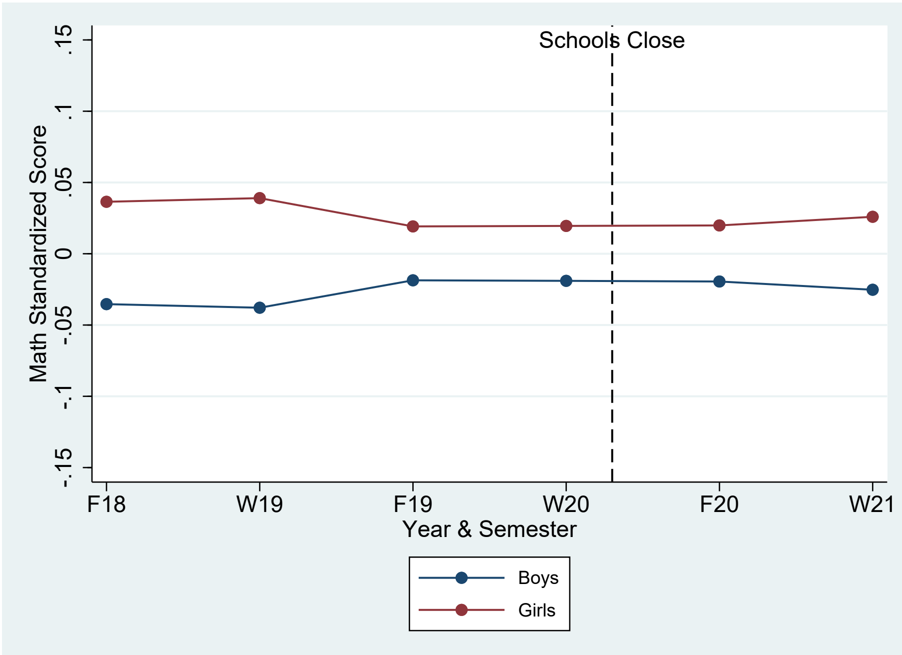


Figure 2: Standardized iReady Assessment Score Trends, SY 2018-2019–SY 2020-2021

(a) Math



(b) Reading

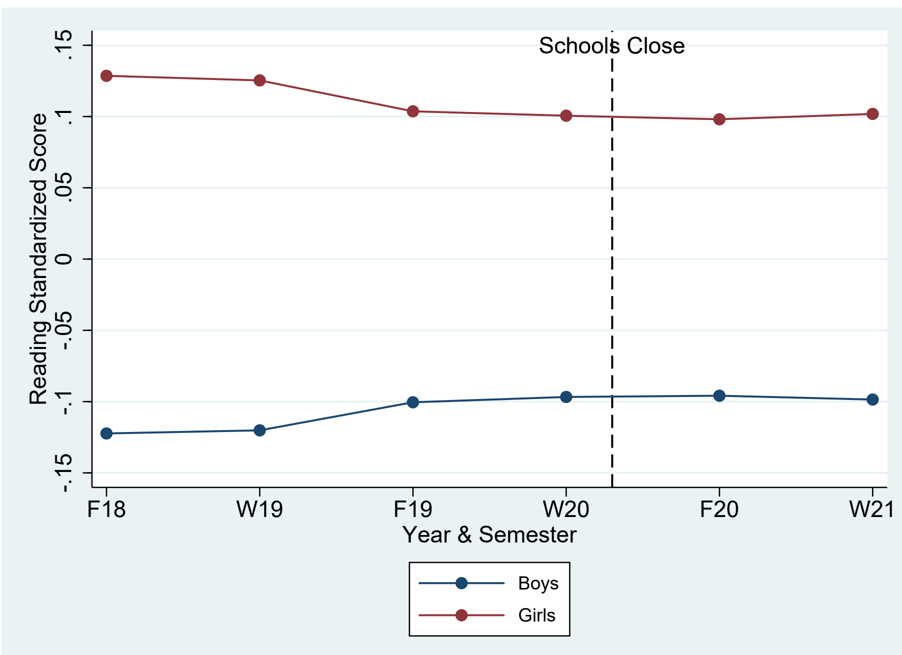
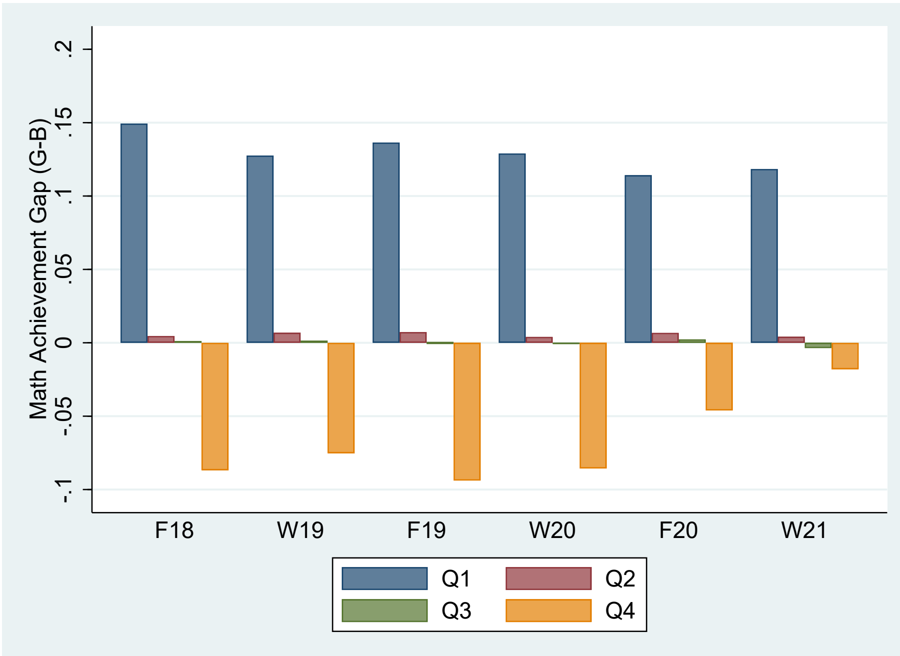


Figure 3: Standardized iReady Assessment Score Gaps (Girls-Boys) by Quartile, SY 2018-2019 – SY 2020-2021

(a) Math



(b) Reading

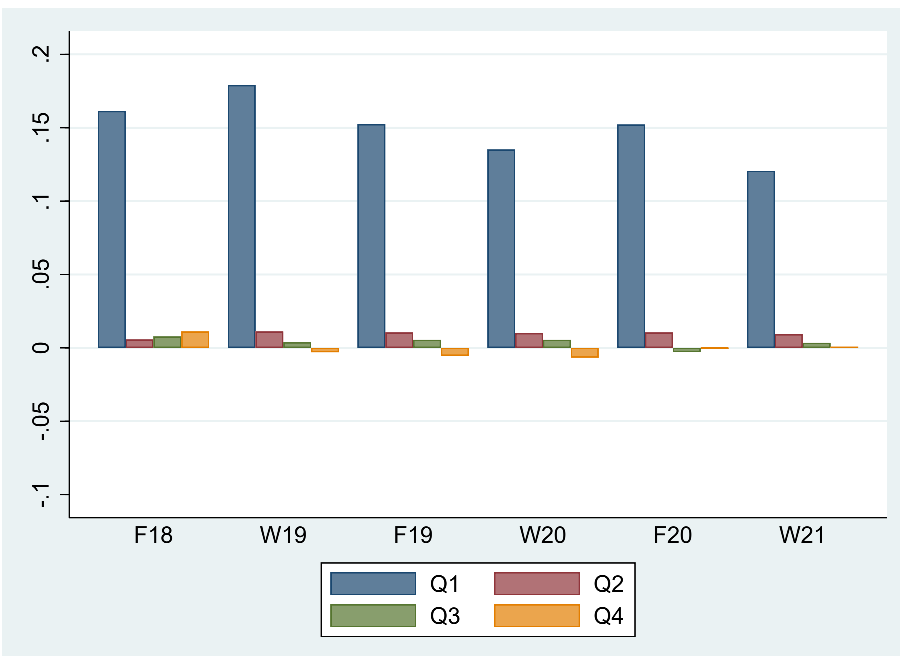
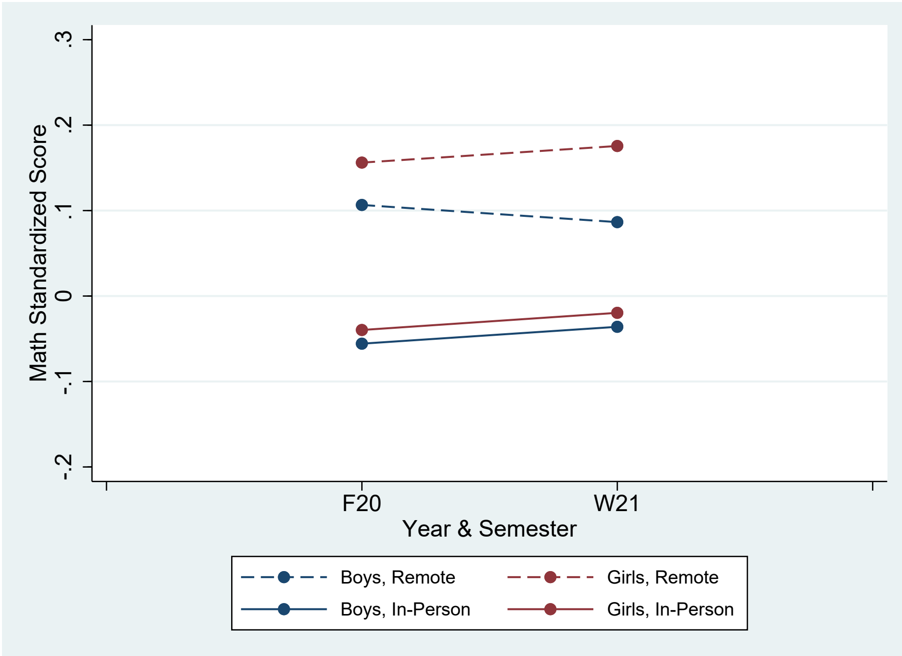
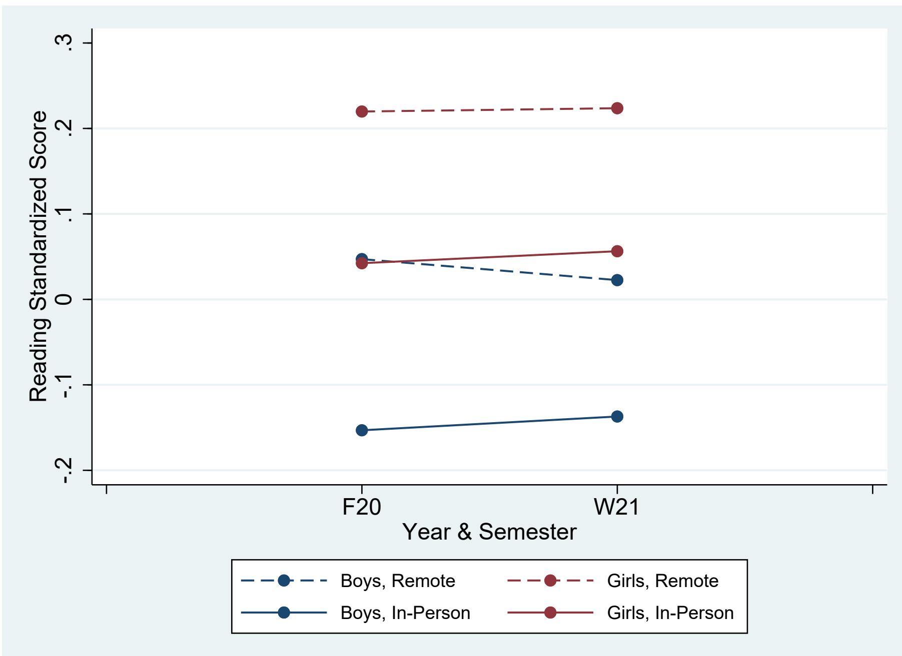


Figure 4: Standardized iReady Assessment Score Trends by Learning Mode, Fall to Winter of SY 2020-2021

(a) Math



(b) Reading



Appendix

A. School Reopening Matrix for the District

Figure A1: School Reopening Matrix for the District

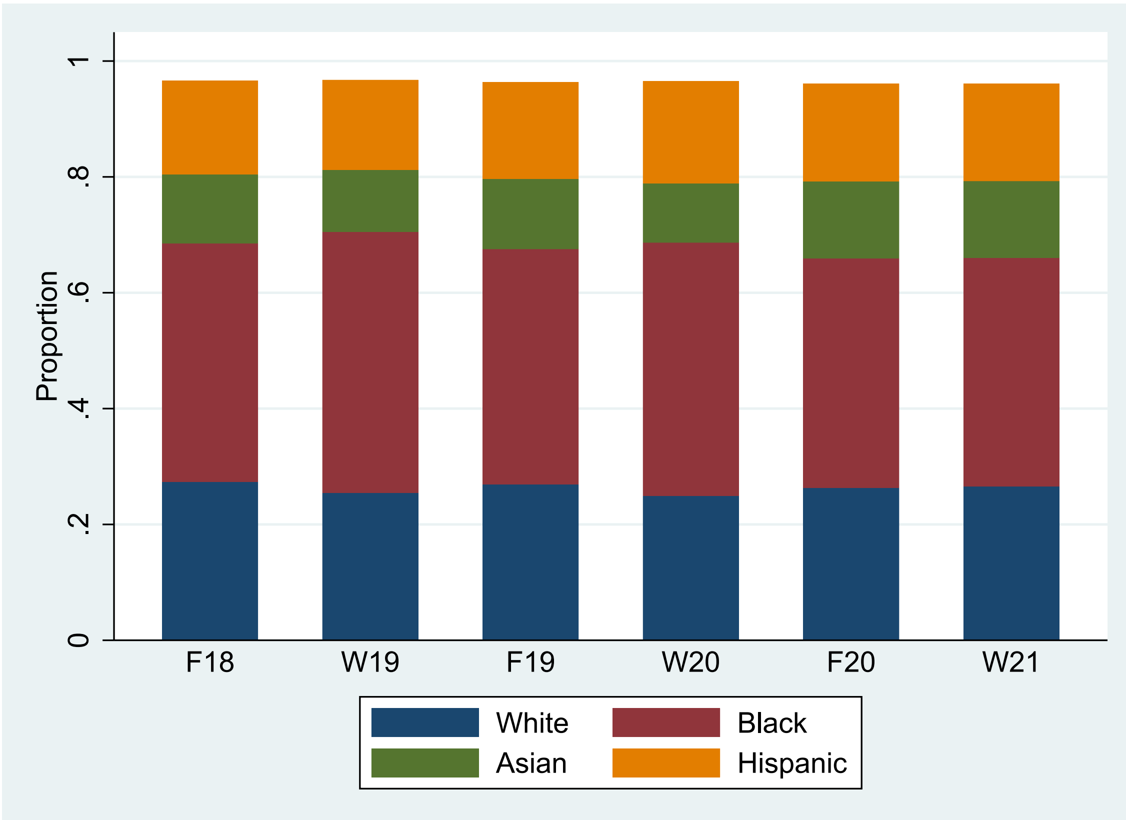
	Universal Remote	Phase I	Phase II	Phase III	Phase IV	Face-to-Face
Criteria to begin Phase		The District plans to move to Phase I of Universal Remote on September 8 to support our students' needs.	The District intends to use the Fulton County Board of Health epidemiology report to determine next steps. The District will begin to transition to the next phase of opening when three consecutive reports show a decline in the New Diagnosis Rate (per last 14 days) of cases per 100,000 OR County-wide New Diagnosis Rate is less than 175 (per last 14 days) per 100,000	The District intends to use the Fulton County Board of Health epidemiology report to determine next steps. The District will begin to transition to the next phase of opening when three consecutive reports show a decline in the New Diagnosis Rate (per last 14 days) of cases per 100,000 OR County-wide New Diagnosis Rate is less than 150 (per last 14 days) per 100,000	The District intends to use the Fulton County Board of Health epidemiology report to determine next steps. The District will begin to transition to the next phase of opening when three consecutive reports show a decline in the New Diagnosis Rate (per last 14 days) of cases per 100,000 OR County-wide New Diagnosis Rate is less than 125 (per last 14 days) per 100,000	The District plans to move to Face-to-Face instruction after the county-wide New Diagnosis Rate is less than 100 per 100,000 cases (per last 14 days)
PreK-2	All remote	90 minutes (1 day per week in 2 sessions)	½ Day (1 day per week)	1 Full Day (1 day per week)	2 Full Days (M/W or T/R)	5 Days
Spec Ed	All remote	180 minutes (1 day per week)	½ Day (1 day per week)	1 Full Day (1 day per week)	2 Full Days (M/W or T/R)	5 Days
3-12	All remote	1:1 by Appointment	½ Day (1 day per week)	1 Full Day (1 day per week)	2 Full Days (M/W or T/R)	5 Days

B. Test Taker Composition

Table B1: The Number of iReady Test Takers by Year and Semester

Grade	Subject	F18/19	W18/19	F19/20	W19/20	F20/21	W20/21
K	Math	5,736	3,614	5,809	5,210	4,177	4,396
	Reading						
1	Math	6,173	4,043	6,031	5,477	5,058	5,221
	Reading						
2	Math	6,267	4,375	6,265	5,771	5,415	5,503
	Reading						
3	Math	6,557	4,685	6,295	5,967	5,784	5,761
	Reading						
4	Math	6,756	4,840	6,589	6,244	5,901	5,862
	Reading						
5	Math	6,752	4,919	6,703	6,308	6,040	5,979
	Reading						
6	Math	6,591	4,789	6,358	4,984	5,594	5,539
	Reading						
7	Math	6,490	5,226	6,624	4,879	5,634	5,489
	Reading						
8	Math	5,699	4,204	6,558	4,244	5,672	5,734
	Reading						

Figure B1: Racial Composition Change of Math and Reading Test Takers, SY 2018-2019–SY 2020-2021



C. Student Disciplinary Incidents

C1. List of Disciplinary Incident Codes

Table C1: List of Disciplinary Incident Codes

Incident	Incident Type	Frequency	Incident	Incident Type	Frequency
0	Continuation of Incident	4,185	22	Weapons – knife[1]	96
1	Alcohol	89	23	Weapons – other[1]	132
2	Arson	15	24	Other Discipline Incident[1]	3,096
3	Battery[1]	3,077	25	Weapons – handgun[1]	17
4	Burglary	61	26	Weapons – rifle[1]	1
5	Computer Trespass	4556	27	Serious Bodily Injury[1]	80
6	Disorderly Conduct[1]	7,964	28	Other firearms	0
7	Drugs, except Alcohol and Tobacco	657	29	Bullying[1]	447
8	Fighting[1]	4,927	30	Other – Attendance Related	3,847
9	Homicide	0	31	Other – Dress Code Violation	48
10	Kidnapping	0	32	Academic Dishonesty	557
11	Larceny or Theft	549	33	Other – Student Incivility[1]	6,141
12	Motor Vehicle Theft	0	34	Other – Possession of Unapproved Items[1]	281
13	Robbery	14	35	Gang-Related[1]	97
14	Sexual Battery[1]	24	36	Repeated Offenses	140
15	Sexual Harassment[1]	221	40	Other Non-Disciplinary Incident	214
16	Sex Offenses[1]	172	42	Electronic Smoking Device[2]	0
17	Threat or Intimidation[1]	1,695	44	Violence Against a Teacher[2]	0
18	Tobacco	727		Total	40,706
19	Trespassing	91			
20	Vandalism	488			

Notes: The table shows a list of disciplinary incident codes and frequency of each incident type during SY 2018-2019 and SY 2019-2020 (incidents prior to the initial school closure on March 18, 2020) from the Student Disciplinary data I use.
[1] : I identify a student as "disruptive" if the student's incident falls into one of these disciplinary incidents.
[2] : These disciplinary incidents were newly listed in GaDOE Discipline Matrix table but none of the students in the analysis sample.

C2. Frequency of Disciplinary Incidents by Student

Table C2: Frequency of Disciplinary Incidents by Student

Number of Incidents	Frequency	Percent	Cumulative Percent
0	12,080	35.35	35.35
1	6,216	18.19	53.55
2	3,846	11.26	64.80
3	2,658	7.78	72.58
4	1,932	5.65	78.23
5	1,482	4.34	82.57
6	1,139	3.33	85.91
7	884	2.59	88.49
8	689	2.02	90.51
9	553	1.62	92.13
9<		7.87	100.00

Note: Table above shows a frequency of disciplinary incidents by sample student during SY 2018-2019 and SY 2019-2020 (incidents prior to the initial school closure on March 18, 2020) from the Student Disciplinary data I use.