GEORGIA
POLICY LABS

Gender Differences in Remote Learning amid COVID-19

Sungmee Kim and Tim R. Sass

Metro Atlanta Policy Lab for Education

November 2023

Background and Motivation

Motivation

Now three years after the COVID-19 pandemic started, a growing body of evidence has fueled concerns over the short- and long-term impacts caused by school closures and unplanned remote learning. A recent release of the National Assessment of Educational Progress (NAEP) long-term trend assessment shows unprecedented declines in reading and math assessment scores from 2020 to 2022, erasing two decades of academic progress in reading and mathematics (National Center for Education Statistics, 2022). While the pandemic dampened average student achievement growth, there is also evidence that pre-existing achievement gaps by race/ethnicity, English language proficiency, and socio-economic status have grown (Skar et al., 2021; Aucejo et al., 2020; Copeland et al., 2021; Bailey et al., 2021; Donnelly & Patrinos, 2021; Dorn et al., 2020; Hammersten et al., 2021; Kuhfeld et al., 2022; Goldhaber et al., 2022). The widening achievement gaps are particularly disconcerting, as they could lead to increased disparities in student outcomes in the long run (Autor et al., 2020; Werner & Woessmann, 2021; Doty et al., 2022).

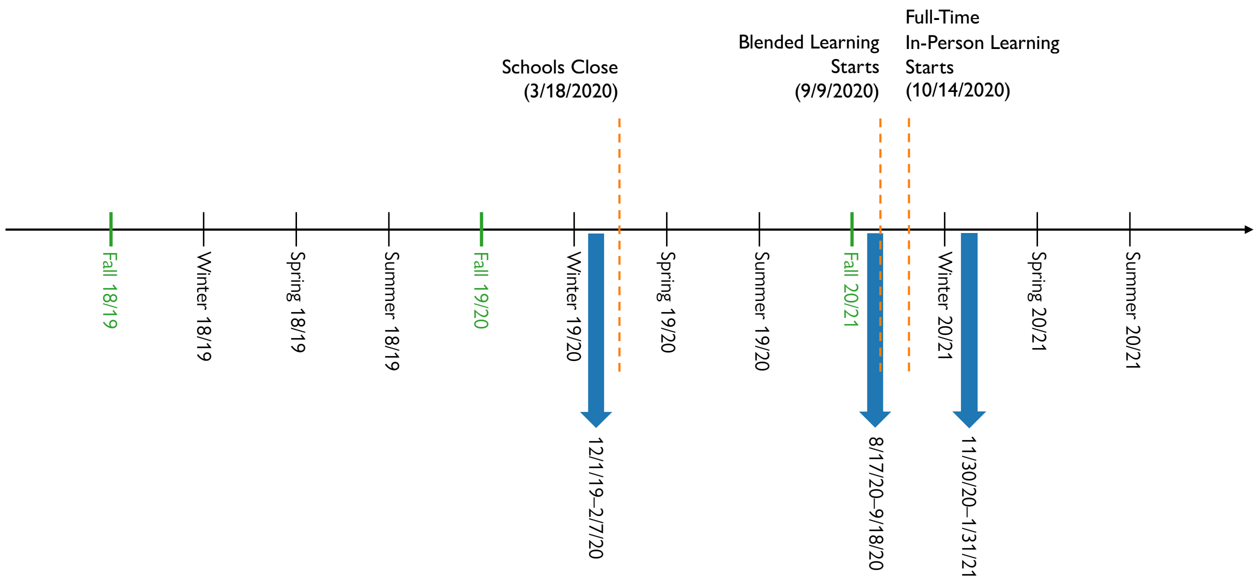
In this report, we explore another dimension of achievement growth differences during the pandemic that has received little attention: student gender. There is emerging evidence of differences in achievement growth for male and female students during the pandemic (Sass & Goldring, 2021), yet there has been little to no work to understand the causes of such gender differences and the implications for education policies. We focus on two potential mechanisms associated with the switch from in-person to remote learning: changes in the nature of peer interactions and differences in the level of self-control required to learn in a remote environment.

Background

Due to the outbreak of COVID-19, Governor Brian Kemp issued an executive order to close all schools in Georgia effective March 31, 2020, and schools remained closed during the remainder of school year (SY) 2019–20 (Georgia Department of Education, 2020). During this roughly nine-week period, schools offered various forms of remote learning to students (Lane, 2020; Sass & Goldring, 2021).

Most school districts in metro Atlanta began SY 2020–21 with fully remote instruction but later offered parents the option of in-person instruction for

Figure 1. Timeline of School Closure and Formative Assessment Testing Windows



their children at varying times in SY 2020–21. The school district we study chose a phased approach for returning to face-to-face instruction. Figure 1 illustrates the phases and timing of the return to full-time in-person instruction in the district.[1] From the start of the school year on August 17 to the beginning of the first phase on September 9, the district provided remote instruction to all students.[2] During Phase I, which lasted less than two weeks, the district gave students in Pre-K through Grade 2 a voluntary opportunity to receive a 90-minute in-person instruction-and-support session once per week. The district provided meals and snacks, and students choosing this option received free transportation.[3] During this initial phase, the district gave students in Grades 3–12 the option to receive support by scheduling one-on-one meetings with their teachers, while continuing their regular remote learning program. Phase I ended on September 21 when the district skipped to Phase III. In Phase III, which only lasted a couple of weeks, the district gave all students the opportunity to attend full-day in-person classes once per week. In Phase IV, which began on October 5, students could attend in-person classes twice a week. Finally, beginning October 14, the district resumed full-day in-person instruction five days per week for all students, though parents could opt to keep their children in full-time virtual learning. Parents had to select a learning mode via an online survey conducted in mid-September. The district enrolled children of parents/guardians who did not respond to the survey (19% of survey recipients) into in-person learning by default.

Table 1. Testing Window for Fall and Winter Exams, SY 2020–21

		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

Table 2. Descriptive Statistics of Number of Attended Days between Fall and Winter Exams, SY 2020–21

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

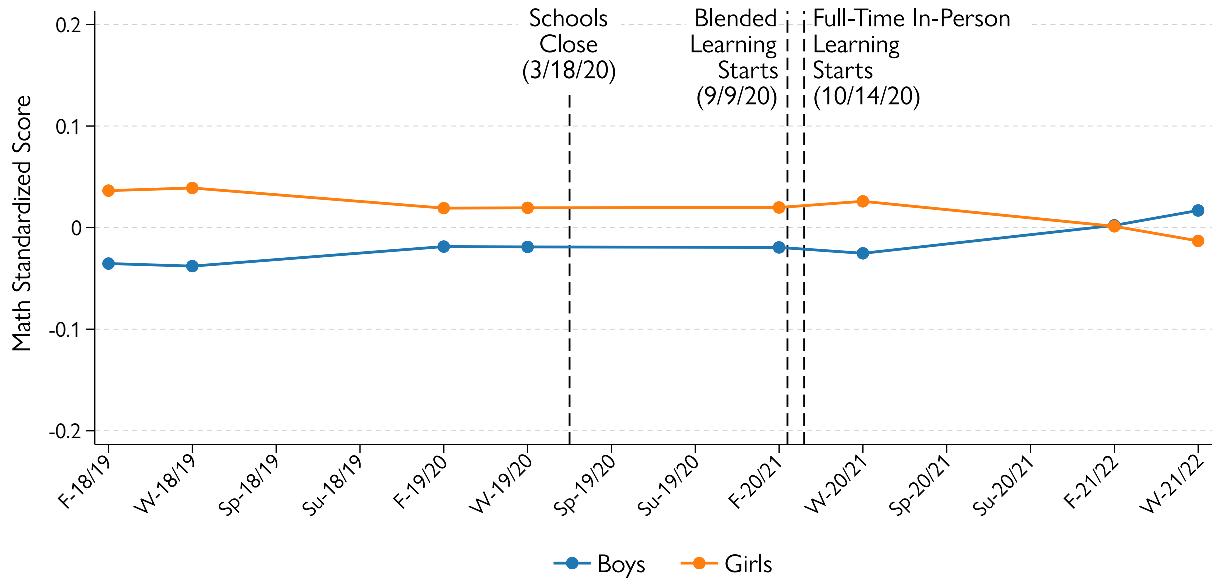
While parents were able to determine student learning mode after October 14, two factors contributed to random variation in student exposure to virtual instruction. First, testing windows for formative assessments are fairly broad, so the dates at which individual students take exams can vary widely (see Tables 1 and 2). Given the phase-in of in-person learning, this translates into differences in exposure to remote learning between assessments. Second, once full-time in-person instruction resumed, the district expected any student who was sick, had a fever, tested positive for COVID-19, or had been exposed to COVID-19 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.

Gender Differences in Student Achievement

Prior research shows that, on average, girls outperform boys on reading/English Language Arts (ELA) exams and either perform similarly to or slightly outperform boys on math exams (Duckworth & Seligman, 2006; Lai, 2010; Sartain et al. (2023); for a meta-analysis: Voyer & Voyer, 2014). We find similar patterns in the metro-Atlanta district we study.

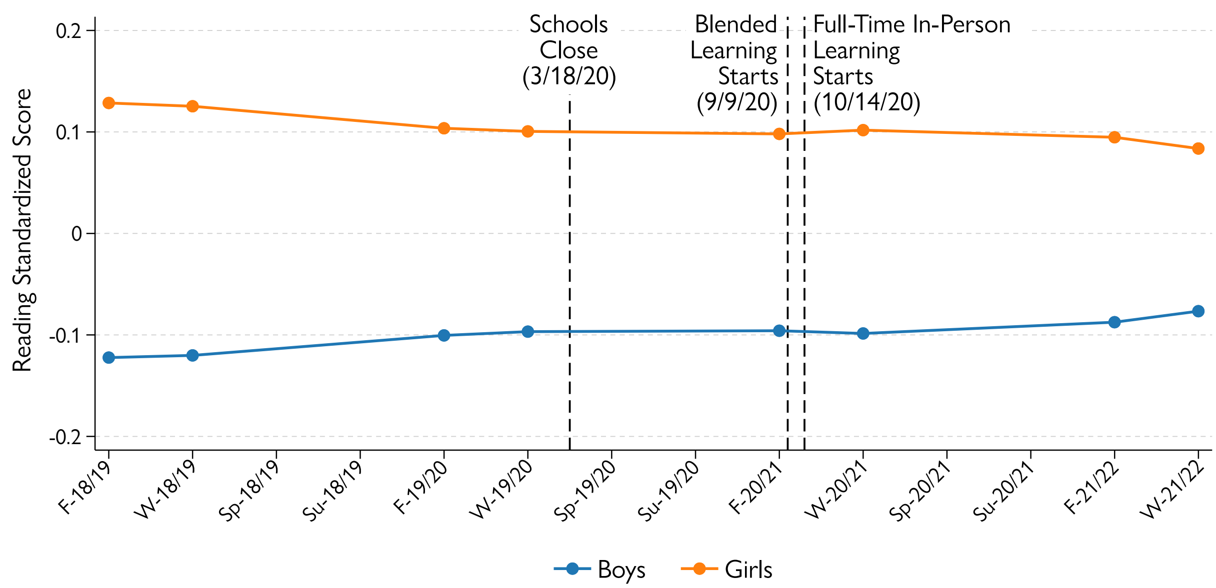
Figures 2A and 2B show trends in standardized math and reading scores trends on the i-Ready formative assessment by gender from SY 2018–19 through SY 2021–22.[4] In the pre-pandemic period (fall SY 2018–19 to winter SY 2019–20), average math scores of girls were consistently higher than the average scores of boys (by as much as 0.1 standard deviations), though the gap declined between winter SY 2018–19 and fall SY 2019–20. The gender achievement gap

Figure 2A. Normalized Math Scores for Students in Grades 1–8, by Gender



Notes. The vertical axis measures the average score in math, where scores are standardized to have a mean of zero and a standard deviation of one across all students within a grade in the district each year/semester.

Figure 2B. Normalized Readings Scores for Students in Grades 1–8, by Gender



Notes. The vertical axis measures the average score in reading, where scores are standardized to have a mean of zero and standard deviation on one across all students within a grade in the district each year/semester.

was considerably larger in reading (0.2 standard deviations or more), though there was also a narrowing of the gap between winter SY 2018–19 and fall SY 2019–20. The gender achievement gaps in math and reading remained relatively constant between the last pre-pandemic exam (winter SY 2019–20) and the first pandemic-era exam in fall (fall 2020-21).

In this report, we focus on the period of transition from fully remote to fully in-person learning: fall SY 2020–21 to winter SY 2020–21. During this period, we observe that average standardized scores for girls rose (relative to the district mean), and average standardized scores for boys declined. These outcomes resulted in an increase in the gender achievement gap, particularly in math.

The Theory behind Gender Achievement Gaps and Instructional Mode

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 & Woessmann, 2021). Compared with traditional face-to-face instruction, students in a virtual learning environment have less direct contact with their peers and teachers because students and teachers are away from physical school buildings and classrooms. Consequently, the pandemic may have led to a temporary reduction in bullying behavior (Bacher-Hicks et al., 2021). Further, self-regulation and parental support and supervision are relatively more important in remote learning (despite challenges for caregivers to do this during the pandemic). Therefore, we could expect that student self-control and self-discipline skills would have a greater impact on student achievement in a remote environment. Similarly, the relative importance of peer inputs (both positive and negative) should decline in remote learning because direct peer interactions are limited.

Prior research finds that innate self-control levels vary by student gender in middle school and could potentially induce differences in achievement gaps between boys and girls (Duckworth & Seligman, 2006; Duckworth et al., 2015). Also, at the post-secondary level, there is evidence that young women respond more to peer influences than do young men (Han & Li, 2009). More generally, there is a growing literature on the role of non-cognitive skills and peers as sources of gender gaps and as factors determining student outcomes (Jacob, 2002; Bertrand & Pan, 2013; Nakajima et al., 2020).

Given differences in the nature of remote instruction vis-à-vis in-person learning and differences between boys and girls in their self-control and susceptibility to peer influences, we consider two potential explanations for growing gender achievement gaps when students spent significant time engaged in remote learning: (a) Remote instruction changed the nature of peer interactions, and girls were less disrupted by their peers' behavior during remote learning. (b) Girls possess better self-control, an essential component of success in remote learning, and learned more than boys did when schools were closed. Understanding the mechanisms driving differential success in pandemic-era remote learning can provide valuable information to properly support students in any future use of remote instruction.

Research Questions

We address the following research questions:

1. Did pandemic-era remote learning dampen any negative effects of having disruptive classmates?
2. To what extent did success in remote learning vary with student self-control?
3. How much of observed gender differences in student outcomes during remote learning can be explained by differences in self-control and exposure to historically disruptive peers?

Data

We combine multiple administrative datasets from a metro-Atlanta school district for SY 2018–19 through SY 2020–21. The resulting student-level panel dataset consists of rich information on student characteristics, such as race/ethnicity, free or reduced-price meals (FRPM) status, English learner (EL) status, types of identified disabilities, and the number and type of disciplinary incidents. The outcome of interest—student achievement growth—is measured by the difference between the scale scores on fall and winter formative assessments in SY 2020–21 for both mathematics and reading.[5] To account for differences in the timing of exam taking, we divide the change in scale scores by the number of instructional days between exams for each student.

Given that testing was limited in high school and involved different exams than those used in elementary and middle school, we initially restrict the analysis to students taking formative assessments in Grades 1–8 during the transitional period (fall and winter tests in SY 2020–21). To determine historical student behavior and its relationship to pre-pandemic student achievement, we further limit the sample to students who were enrolled in the district in either SY 2018–19 or SY 2019–20 and took at least one pre-pandemic formative assessment in Grades 1–8 during that time. The requirement of having at least one prior score limits the analysis of student achievement growth in the transitional period to students in Grades 2–8.

The two independent variables of primary interest are the proportion of disruptive classmates and a student's self-control level. The proportion of historically disruptive peers in a classroom is constructed by linking Student Class and Student Discipline data. The Student Class file includes information on which classes students took in each semester, and the Student Discipline file provides student incident-level information on the type and intensity of each disciplinary incident. A student is considered "historically disruptive" if the student committed any potentially disruptive disciplinary incidents any time from the start of SY 2018–19 up to the time of school closures in mid-March 2020. Potentially disruptive incidents include bullying, fighting, sexual battery, sexual harassment, sex offenses, threat or intimidation, carrying weapons (e.g., knife, handgun, rifle) and other firearms, serious bodily injury, disorderly conduct, and student incivility.[6] To measure the extent of disruptive peers, we calculate the proportion of students in each math and reading/ELA class that are designated as historically disruptive. We measure individual exposure to historically disruptive peers by the classroom proportion of historically disruptive students averaged over the math or reading classes a student was enrolled in during the relevant time period.[7]

The second key variable of interest—student's self-control level—is proxied by spending little time answering questions on exams.[8] The formative assessment results contain a "rush flag" indicator, signifying a student's average time on each task of the exam was shorter than a designated length of time.[9] We measure (lack of) self-control with an indicator which equals 1 if a student ever rushed on a formative exam at any time in the pre-pandemic sample period.

To gauge student exposure to remote learning, we employ individual-level daily blending learning attendance data collected by the district from the beginning of SY 2020–21 through January 2021.[10] The data cover the period in which students transitioned from universal remote instruction to in-person instruction and span the time between formative assessments given in fall and winter of SY 2020–21. The blended learning attendance data indicate the assigned learning mode for each student on each instructional day and whether the student attended that day.[11] Combining these data with the dates each student took the fall and winter assessments allows us to determine the proportion of days attended between exams in which each student attended remote instruction. As students take exams over a period of several weeks each semester, two students who switched from remote to in-person instruction on the same day can still have different proportions of remote learning between exams.

Lastly, we use the district's Parental Survey data, which contains student-level information on parents' preferences for instructional mode and types of transportation to/from school in SY 2020–21, and school-level information on the number of COVID-19 positive and quarantined cases to predict the proportion of days a child experiences remote learning.

Methodology

To measure the relationships between peer behavior and self-control on student learning, we estimated multivariate regression models of student achievement that include controls for student demographics and school-level factors that may affect student achievement.[12] The estimated coefficients for peer composition and for own self-control thus represent the partial correlations between these variables and student achievement (holding other factors constant).

We estimate the regression models of average daily achievement gains over two distinct periods: (a) a pre-pandemic period covering fall to winter of SY 2018–19 and fall to winter of SY 2019–20 and (b) the period of transition from planned remote learning to in-person instruction (fall to winter of SY 2020–21). The pre-pandemic period analysis measures the effects of peer composition and own self-control in a typical in-person learning environment. For the transition period, we interact the proportion of time spent in remote instruction with the variables representing exposure to disruptive peers and self-control to gauge how remote learning affected the relationships between self-control and student achievement and the relationships between exposure to disruptive peers and student achievement.

Because parental choice partly determines exposure to remote learning, the exposure measure could reflect unobserved factors (e.g., parental resources) that affect both learning mode and student achievement. As discussed in the Appendix, we employ an alternative estimation procedure, known as a two-stage-least squares (2SLS) regression model, to address potential bias from self-selection. For the standard multivariate regression model and the 2SLS model,

Table 3. Pre-pandemic Summary Statistics, Full Sample and by Gender

	Full sample		Girls		Boys		Mean difference (G-B)
	Mean	SD	Mean	SD	Mean	SD	
Scale-score growth per day (math)	0.171	0.237	0.166	0.218	0.176	0.254	-0.010***
Scale-score growth per day (reading)	0.232	0.398	0.229	0.369	0.236	0.425	-0.008**
Committed a disruptive incident	0.063	0.242	0.031	0.172	0.094	0.292	-0.063***
Proportion of disruptive peers (math)	0.068	0.104	0.066	0.100	0.070	0.109	-0.004***
Proportion of disruptive peers (reading)	0.073	0.106	0.070	0.102	0.076	0.110	-0.006***
Ever rushed (math)	0.179	0.383	0.131	0.338	0.225	0.418	-0.094***
Ever rushed (reading)	0.130	0.336	0.089	0.285	0.169	0.374	-0.079***
N (math)	53,388		26,375		27,013		
N (reading)	48,651		23,898		24,753		
Test takers (math)	36,091		17,817		18,280		
Test takers (reading)	35,593		17,492		18,073		

Notes. Analyses sample includes students in Grade 1 to Grade 7 enrolled in public schools located in the school district during the pre-pandemic semesters (fall and winter of SY 2018–19 and 2019–20 but prior to the initial school closure). The unit of the number of observations is individual in each school-year-semester. *: mean difference is statistically significant at the 10% level; **: mean difference is statistically significant at the 5% level; ***: mean difference is statistically significant at the 1% level

we decompose the gender achievement differential into two components: the part that is "explained" by observable differences between boys and girls and the part that is due to differences in unobserved characteristics of boys and girls. Within the explained part, we decompose differences in test scores between boys and girls due to differences in the levels of observed characteristics (including self-control and the proportion of historically disruptive peers) and the proportion due to differences in the marginal impacts of changes in student characteristics (including self-control and exposure to disruptive peers).

Finding 1: Pre-pandemic Gender Differences

Prior to the pandemic, boys were more likely than girls to exhibit disruptive behavior: 3% of girls and 9% of boys had a disruptive disciplinary incident. Girls exhibited more self-control: Boys were 1.7 times more likely to ever rush on math exams and 1.9 times more likely to ever rush on reading exams.

Table 3 provides summary statistics for students in Grades 1–7 for the beginning of each of the two semesters prior to the pandemic: between fall and winter exams of SY 2018–19 and SY 2019–20.[13] In math and reading, on average, 7% of classmates have a history of disruptive behavior. **Boys were much more likely to have a history of disruptive behavior.** 9% of boys have been disciplined for disruptive behaviors, while only 3% of girls had a record of one or more prior incidents of disruptive behavior.

Similarly, **girls exhibited greater self-control than boys.** On average, 13% of girls were flagged for ever rushing on math formative assessments, while 23% of boys ever rushed through math exams. In reading, 9% of girls and 17% of boys ever rushed through the formative assessments during the pre-pandemic semesters.

Finding 2: Pre-pandemic Effects of Disruptive Peers and Lack of Self-Control on Student Achievement

> Pre-pandemic, a doubling of the proportion of disruptive peers in a class (from the mean of 7% to 14%) is associated with a decrease in math achievement growth per day of about 2% pre-pandemic. Having previously rushed through an exam is associated with a reduction in reading score growth of about 13%.

Table 4 shows estimates of the determinants of student achievement growth during the pre-pandemic period. Separate estimates are provided for the full sample and by gender. Consistent with expectations, we find that increases in the proportion of disruptive peers and having previously rushed through an exam have a negative impact on student achievement growth for boys in math and in reading, respectively. However, we did not find negative effects for girls in either subject. Across all students, we find that doubling the proportion of historically disruptive peers in classrooms from the mean of 7% to 14% is associated with a reduction in the daily growth in math formative assessment scores of 0.004 scale-score points or about 2% of the average growth of 0.171 scale-score points per instructional day (shown in Table 3). Being an "ever-rusher" is associated with a decrease in reading-test-score growth of 0.030

Table 4. Ordinary Least Squares Regression Estimates of the Determinants of Student Achievement Growth per Instructional Day by Subject, Pre-pandemic Period

	Full sample		Girls		Boys	
	Math	Reading	Math	Reading	Math	Reading
Proportion of disruptive peers	-0.050**	-0.012	-0.031	-0.026	-0.047*	0.010
	(0.020)	(0.037)	(0.021)	(0.035)	(0.025)	(0.043)
Ever rushed	0.008*	-0.030***	0.006	-0.002	0.008	-0.027***
	(0.005)	(0.010)	(0.006)	(0.013)	(0.005)	(0.010)
Female	-0.010***	-0.011***				
	(0.002)	(0.004)				
Female × proportion of disruptive peers	0.023	0.008				
	(0.024)	(0.043)				
Female × ever rushed	-0.004	0.031*				
	(0.008)	(0.017)				
Black (ref. group = White)	-0.009**	-0.011	-0.009*	-0.011	-0.009*	-0.011
	0.004)	(0.007)	(0.005)	(0.009)	(0.006)	(0.010)
Asian	-0.001	-0.014**	0.002	-0.015*	-0.004	-0.013
	(0.003)	(0.006)	(0.004)	(0.008)	(0.005)	(0.009)
Hispanic	-0.006*	-0.012*	-0.008	-0.023***	-0.004	-0.000
	(0.004)	(0.006)	(0.005)	(0.009)	(0.006)	(0.009)
Other non-White	-0.007	-0.018*	-0.006	-0.009	-0.006	-0.028*
	(0.005)	(0.010)	(0.007)	(0.013)	(0.008)	(0.014)
FRPM	-0.013***	0.002	-0.011***	0.004	-0.015***	-0.001
	(0.003)	(0.005)	(0.004)	(0.007)	(0.004)	(0.007)
EL	0.020***	0.061***	0.026***	0.078***	0.014**	0.046***
	(0.004)	(0.008)	(0.006)	(0.011)	(0.006)	(0.012)
Identified disability status	-0.001	0.006	-0.005	0.009	0.000	0.004
	(0.004)	(0.007)	(0.008)	(0.011)	(0.005)	(0.009)
Number of incidents	-0.008***	-0.001	-0.002	-0.010	-0.011***	0.003
	(0.003)	(0.008)	(0.004)	(0.010)	(0.004)	(0.010)
Grade FE	Y		Y		Y	
School FE	Y		Y		Y	
Year-semester FE	Y		Y		Y	
N	53,388	48,651	26,375	23,898	27,013	24,753

Notes. Analyses sample includes students in Grade 1 to Grade 7 enrolled in public schools located in the school district during the pre-pandemic semesters (fall and winter of SY 2018–19 and 2019–20 but prior to the initial school closure. Robust standard errors in parentheses below estimated coefficients. The unit of observation is the individual in each school-year-semester; if a student was observed throughout the pre-pandemic semesters, there would be two observations for each student. Outcome variables are standardized math and reading achievement scores. *: coefficient is statistically significant at the 10% level; **: coefficient is statistically significant at the 5% level; ***: coefficient is statistically significant at the 1% level scale score points or about 13% of the average daily growth rate in reading of 0.232 scale-score points.

Finding 3: Determinants of Gender Differences in Achievement Growth During Remote Learning

For boys, student achievement growth diminishes the greater the proportion of time spent in remote learning, lowering achievement growth by 27–35%. However, for girls, the impacts of learning mode on achievement growth are modest, ranging from -6% to +2%. Increases in the proportion of time spent in remote instruction tends to mitigate any negative effects of historically disruptive peers in math for girls. However, we do not find statistically significant effects for the interaction of remote learning and historically disruptive peers in reading for either girls or boys. Lack of self-control (proxied by rushing on pre-pandemic exams) does not appear to have a substantial effect on learning gains in remote instruction for either boys or girls.

To gauge how gender achievement gaps changed during remote learning and the mechanisms for those changes, we estimate a multivariate regression model of student achievement growth per instructional day between the fall and winter formative assessments in SY 2020–21. We allow the effects of historically disruptive peers and self-control (proxied by rushing on prior exams) to vary by learning mode and by gender. We present coefficient estimates from both an ordinary least squares (OLS) and a two-stage least squares (2SLS) model in Table 5. Estimates from the 2SLS model are similar in magnitude to those from the OLS model, though are somewhat less precise (particularly in math).

Based on the OLS model, we find that **remote instruction substantially reduces student achievement growth for boys but has only a relatively small effect for girls.** All else equal, a boy who spent all the time between exams in remote learning would learn 0.047 scale score points per day less in math and 0.082 scale-score points less in reading per day than a boy who received

Table 5. Ordinary and Two-Stage Least Squares Regression Estimates of the Determinants of Student Achievement Growth per Instructional Day by Subject, Transitional Period

	OLS		2SLS	
	Math	Reading	Math	Reading
Proportion of disruptive peers	-0.016 (0.059)	0.050 (0.104)	0.014 (0.071)	0.007 (0.131)
Ever rushed	0.006 (0.019)	0.035 (0.037)	0.018 (0.021)	0.014 (0.045)
Proportion of remote days	-0.047*** (0.013)	-0.082*** (0.019)	-0.025 (0.015)	-0.077*** (0.022)
Proportion of remote days × proportion of disruptive peers	-0.127 (0.088)	-0.007 (0.151)	-0.181 (0.111)	0.036 (0.192)
Proportion of remote days × ever rushed	0.004 (0.029)	-0.027 (0.057)	-0.015 (0.035)	0.007 (0.071)
Female	-0.026** (0.010)	-0.023 (0.015)	-0.009 (0.012)	-0.015 (0.017)
Female × proportion of disruptive peers	-0.008 (0.081)	0.107 (0.142)	-0.023 (0.100)	0.087 (0.178)
Female × ever rushed	0.017 (0.028)	-0.046 (0.061)	0.008 (0.032)	0.013 (0.072)
Female × proportion of remote days	0.051*** (0.016)	0.069*** (0.023)	0.025 (0.018)	0.066** (0.027)
Female × proportion of remote days × proportion of disruptive peers	0.218* (0.123)	-0.175 (0.209)	0.250 (0.155)	-0.215 (0.263)
Female × Proportion of remote days × ever rushed	-0.061 (0.043)	0.036 (0.093)	-0.048 (0.053)	-0.068 (0.114)
Black (ref. White)	-0.014* (0.008)	-0.011 (0.012)	-0.020*** (0.007)	-0.037*** (0.011)
Asian	0.011 (0.007)	0.032*** (0.010)	0.007 (0.007)	0.045*** (0.010)
Hispanic	-0.001 (0.008)	0.009 (0.011)	0.006 (0.008)	0.005 (0.012)
Other non-White	0.005 (0.012)	-0.022 (0.019)	0.011 (0.012)	-0.019 (0.020)
FRPM	-0.019*** (0.007)	-0.020* (0.011)	-0.028*** (0.006)	-0.052*** (0.010)
EL	-0.006 (0.011)	0.034** (0.015)	-0.015 (0.012)	0.034** (0.016)
Identified disability status	-0.029*** (0.008)	-0.043*** (0.013)	-0.025*** (0.009)	-0.036*** (0.014)
Number of disciplinary incidents	-0.049 (0.032)	0.000 (0.044)	-0.049 (0.035)	-0.002 (0.044)

Table 5. Ordinary and Two-Stage Least Squares Regression Estimates of the Determinants of Student Achievement Growth per Instructional Day by Subject, Transitional Period

	OLS		2SLS	
	Math	Reading	Math	Reading
Grade FE	Y		Y	
School FE	Y		omitted	
Year-Semester FE	omitted		omitted	
N	25,547	29,340	23,297	26,520

Notes. Sample includes students in Grade 2 to Grade 8 enrolled in public schools located in the school district during the transitional period (fall and winter of SY 2020–21). Robust standard errors in parentheses below estimated coefficients. The unit of the number of observations is individual in each school-year-semester. Outcome variables are standardized math and reading achievement scores. *: coefficient is statistically significant at the 10% level; **: coefficient is statistically significant at the 5% level; ***: coefficient is statistically significant at the 1% level

all instruction in-person. Based on pre-pandemic achievement growth rates, these impacts are equivalent to a 27% reduction in math and 35% reduction in reading. For girls, the impact of full-time remote instruction in math is 0.004 scale-score points per day or a two-percent increase in achievement growth relative to the pre-pandemic mean for girls. The impact of full-time remote instruction in reading is -0.013 scale-score points per day or a six-percent reduction in achievement growth relative to pre-pandemic norms.

We do not find strong effects of personal self-control (as measured by prior rushing on exams) on achievement growth during the transition from remote to in-person learning. In math and reading, we cannot confidently rule out that the effect of prior rushing on daily average achievement growth is zero. Likewise, **we do not find that prior rushing behavior alters the magnitude of the impact of remote learning on student achievement growth. This is true for boys and girls.**

We also find relatively little impact of peer behavior on achievement growth per day during the transitional period. With completely in-person instruction (i.e., when proportion of remote days is zero), increases in the proportion of historically disruptive peers is not significantly correlated with daily achievement growth in either math or reading for boys or for girls. For boys, the lack of a relationship between the proportion of historically disruptive peers and achievement growth in either subject does not change as the proportion of time spent in remote instruction increases. **Contrary to expectations, we find a marginally significant increase in the rate of achievement growth per day in math for girls as the proportion of disruptive peers increases when**

instruction is fully remote, but the effect is small. A doubling of the proportion of disruptive peers from the pre-pandemic mean of 7% to 14% would result in a 0.005 scale-score-point increase in achievement growth in math. This is equivalent to a three-percent increase relative to the pre-pandemic average for girls.

As discussed above, the choice of learning mode (once in-person instruction was available) could be endogenous, meaning that unobserved student and family characteristics could be influencing both the proportion of days in remote instruction and student achievement. This would lead to biased estimates of the effect of remote learning as well as the effect of any factors that are interacted with the proportion of days in remote instruction, such as (Proportion of Remote Days x Proportion of Disruptive Peers). The 2SLS model estimates presented in columns 3 and 4 of Table 5 should be free of such bias, but they may be less precise because the 2SLS approach replaces the actual exposure to remote learning with a predicted value based on a set of exogenous instruments described above. The 2SLS estimates are qualitatively similar to the OLS estimates but, as expected, are generally less precise.

Finding 4: The Relative Contributions of Self-Control and Reduced Exposure to Disruptive Peers in Explaining Gender Differences in Achievement Growth During Remote Learning

> By far, the largest contributor to gender differences in achievement growth during the pandemic was the superior achievement growth of girls while in remote instruction. We find conflicting evidence for causes of girls' relative success in remote instruction, however. For students in math classes with a substantial share of historically disruptive students, the advantage of girls in remote learning was even higher, while having a history of rushing through exams lessened the learning advantage of girls in remote instruction. In contrast, the effects of having disruptive peers and lack of self-control had just the opposite effect on the relative performance in reading of girls in remote learning.

Table 6. Mean Student Characteristics for Decomposition Calculation by Gender

	Math			Reading		
	Girls	Boys	G-B	Girls	Boys	G-B
Achievement growth per day (fall to winter SY 2020–21)	0.118	0.107	0.011	0.129	0.115	0.014
Proportion of remote days	0.581	0.566	0.015	0.586	0.570	0.016
Proportion of disruptive peers	0.072	0.075	-0.003	0.070	0.078	-0.008
Ever rushed	0.118	0.200	-0.082	0.071	0.135	-0.064
Proportion of remote days × proportion of disruptive peers	0.045	0.044	0.001	0.043	0.047	-0.004
Proportion of remote days × ever rushed	0.068	0.110	-0.042	0.042	0.076	-0.034
Black	0.380	0.372	0.008	0.372	0.368	0.004
Asian	0.139	0.139	0.001	0.141	0.137	0.004
Hispanic	0.164	0.171	-0.007	0.170	0.173	-0.003
Other non-White	0.038	0.034	0.004	0.038	0.035	0.003
FRPM	0.428	0.429	-0.001	0.433	0.428	0.005
EL	0.070	0.086	-0.016	0.074	0.090	-0.016
Identified disability status	0.076	0.145	-0.069	0.083	0.153	-0.070
N	12,643	12,904		14,376	14,964	

Notes. Sample includes students in Grade 2 to Grade 8 enrolled in public schools located in the school district during the transitional period (fall and winter of SY 2020–21). The unit of the number of observations is individual in each school-year-semester. Gender differences are computed by first rounding the gender-specific means to the nearest 0.001. The unrounded means are used for the decomposition calculation presented in Table 7. *mean difference is statistically significant at the 10% level; **: mean difference is statistically significant at the 5% level; ***: mean difference is statistically significant at the 1% level

Above, we showed that, prior to the pandemic, boys experienced higher achievement growth per day than did girls in math and reading (see Table 3). However, when a large part of instruction was remote (fall to winter of SY 2020–21), the gender difference flipped, and girls experienced greater achievement growth—particularly in math (see Table 6). Here, we decompose the gender difference in achievement growth during the transitional period in order to determine the extent to which remote instruction was driving the difference in achievement growth between girls and boys and the extent to which the differential effects of remote learning depended peer influences and own self-control.

We divide the gender difference in achievement growth during the transition period into two components: (a) the part that is due to differences in the average characteristics of boys and girls and (b) the part that is due to differences in the impact of remote learning and how that impact varies with the extent of exposure to historically disruptive peers and self-control. Within the category of differences in average characteristics, we differentiate between observable characteristics like race and identified disability status, and unobserved traits, such as student motivation and parental support.

We provide the differences in the characteristics of boys and girls for the sample used to compute the OLS estimates of the determinants of achievement growth in Table 6.[14] We then use those differences in average characteristics and the estimated impacts of characteristics on student achievement growth reported in Table 5 to determine the proportions of the remote learning gender achievement gap that can be explained by differences in characteristics between boys and girls and the payoff to those characteristics.[15]

Table 7 reports the results from decomposing the gender difference in achievement growth during the transitional period for math and for reading, respectively, based on the OLS model.[16] Focusing on the results for math, two components stand out. First, **differences in the unobserved characteristics of girls and boys produce an advantage in achievement growth for boys that is about two and one-half times (249%) the advantage in achievement growth experienced by girls in the transition period. Full-time remote instruction more than completely cancels out that deficit**, with an impact on the differential in achievement growth equal to 277% of the difference in achievement growth per day between girls and boys in the transition period.

About 22% of the superior math achievement growth of girls during the transition period is due to differences in observable student characteristics. The majority of the demographic difference is due to the lower incidence of identified disabilities among girls. The remainder of differences in average observable characteristics are relatively modest, ranging from -6.8% to +0.5% of the observed gender gap in student achievement growth during the transition period.

Of primary interest are the differences in achievement growth associated with increases in remote instruction and how the effect of increased remote instruction varies with the proportion of disruptive peers and self-control. For example, consider students who spent 60% of days attended between math exams in remote instruction (roughly the average amount) but had no disruptive peers and who never rushed through exams in the pre-pandemic period. **Based on gender differences in the impact of remote instruction, the effect of 60% of instructional days in remote instruction would explain 166%**

Table 7. Decomposition of the Achievement Growth Per Day Difference between Girls and Boys (OLS Estimates), Transitional Period

	Math		Reading	
	Amount	% of Total	Amount	% of Total
Total Gender Achievement Growth Gap	0.011	100.0	0.014	100.0
Total Gap Due to:				
Mean Difference in:				
Unobserved Student Characteristics	-0.026	-249.3	-0.023	-165.9
Observed Student Demographics	0.002	22.0	0.002	16.7
School/Grade Enrolled In	-0.000	-2.2	-0.001	-6.6
Proportion of Disruptive Peers	0.000	0.5	0.000	-3.1
Ever Rushed	-0.001	-5.0	-0.002	-16.1
Proportion of Remote Days	-0.001	-6.8	-0.001	-9.7
Proportion of Remote Days × Proportion of Disruptive Peers	0.000	-0.3	0.000	0.2
Proportion of Remote Days ×Ever Rushed	0.000	-1.5	0.001	6.5
Impact Difference in:				
Proportion of Disruptive Peers	-0.001	-5.8	0.007	53.6
Ever Rushed	0.002	18.5	-0.003	-23.4
Proportion of Remote Days	0.029	276.8	0.041	291.5
Proportion of Remote Days × Proportion of Disruptive Peers	0.010	91.7	-0.008	-54.7
Proportion of Remote Days × Ever Rushed	-0.004	-38.6	0.002	10.9

Notes. Sample includes students Grade 2 to Grade 8 enrolled in public schools located in the school district during the transitional period (fall and winter of SY 2020-21). The unit of the number of observations is individual in each school-year-semester. Decomposition calculation is based on OLS estimates from Table 5, columns 1 and 3, and mean statistics of girls and boys in the analysis sample.

of the transition-period gender math achievement gap (0.6 x 276.8). For students who were "ever rushers" and spent 60% of days in remote instruction, the difference in the impact of remote learning would explain 143% or ((0.6 x 276.8) + (0.6 x -38.6)) of the gap. For non-rushers, having 7% (historically) disruptive classmates (the pre-pandemic average) increases the proportion of the gender gap in the transition period explained by the impact of remote learning from 166% to 172% or ((0.6 x 276.8) + (0.07 x 91.7)).

For reading, the magnitude and direction of the impact of remote learning on achievement growth is similar to those in math, explaining 291.5% of the gender achievement growth gap. However, the effects of being a “rusher” and having disruptive peers have the opposite effects on the gender achievement gap than they do in math.

Discussion

We find substantial differences between boys and girls in their achievement growth during the transition from remote learning back to in-person instruction. Remote learning substantially reduces student achievement growth for boys but has only a relatively small effect for girls. We posited two mechanisms for the relatively strong performance of girls in remote learning: reduced exposure to disruptive peers and better self-control. Consistent with our expectations, we find that the advantage girls have in achievement growth during remote learning in math increases with the proportion of historically disruptive peers. However, we cannot say with confidence that gender differences in remote learning in reading vary with peer-group composition. In both math and reading, we cannot confidently say that self-control (as measured by prior rushing on exams) alters the advantage of girls in remote learning in either math or reading. Correspondingly, while the share of the gender achievement gap during the transition back to in-person learning that is attributable to remote instruction is large, the advantage girls have in remote learning does not systematically vary with peer composition and student self-control.

Our findings suggest that there was considerable variation between girls and boys in their ability to successfully navigate remote learning. However, it does not appear that much of the difference can be explained by our (somewhat crude) measure of self-control, and a large proportion is unexplained by observable differences between boys and girls. At a minimum, this suggests that more work needs to be done to identify students who struggle in remote learning and target supports to those who are struggling.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A230400 to Georgia State University Research Foundation, Inc. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

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, 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. (2021). *The COVID-19 pandemic disrupted both school bullying and cyberbullying.* NBER Working Paper 29590.

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

Bertrand, M. and J. Pan (2013). *The trouble with boys: Social influences and the gender gap in disruptive behavior.* American *Economic Journal: Applied Economics*, *5*(1): 32–64.

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.

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, December, 8, 224–228.

Doty, E., Kane, T. J., Patterson, T., & Staiger, D. O. (2022). *What Do Changes in State Test Scores Imply for Later Life Outcomes?* NBER Working Paper 30701.

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.

Georgia Department of Education. (2020, August 1). *COVID-19 (Coronavirus) and Schools.* Georgia Insights of GaDOE. 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.* NBER Working Paper 30010.

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

Hammerstein, S., König, C., Dreisörner, T. and Frey, A. (2021). *Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review.* Front. Psychol. 12:746289.

Hitt, C., Trivitt, J., & Cheng, A. (2016). *When you say nothing at all: The predictive power of student effort on surveys. Economics of Education Review*, *52*, 105–119.

Hitt, C. E. (2015). *Just filling in the bubbles: Using careless answer patterns on surveys as a proxy measure of noncognitive skills.* University of Arkansas Department of Education Reform Working Paper, 6.

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.

Kuhfeld, M., Soland, J., & Lewis, K. (2022). *Test score patterns across three COVID-19-impacted school years.* Annenberg Institute Working Paper 22-521.

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. gpb.org/blogs/education-matters/2020/03/16/gov-kemp-orders-all-k-12-georgia-schools-close-until-end-of

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.* nationsreportcard.gov/highlights/ltt/2022/

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). Retrieved from Annenberg Institute at Brown University: doi.org/10.26300/b0py-tz14

Sass, T., & Goldring, T. (2021). *Student Achievement Growth During the COVID-19 Pandemic*. Georgia Policy Labs. gpl.gsu.edu/publications/student-achievement-growth-during-the-covid-19-pandemic/

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*.

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

Werner, K., & Woessmann, L. (2021). *The legacy of covid-19 in education*. EdWorkingPaper: 21-478.

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.

Zamarro, G., Cheng, A., Shakeel, M. D., & Hitt, C. (2018). *Comparing and validating measures of non-cognitive traits: Performance task measures and self-reports from a nationally representative internet panel. Journal of Behavioral and Experimental Economics*, *72*, 51–60.

Endnotes

1. Details of the district's school reopening plan are provided in Appendix Table A1.

2. Students in Grades 2 through 12 received a device issued by the school district.

3. The district provided transportation for students in Grades 3–5 and all middle and high school students returning for face-to-face instruction during future phases.

4. Scores have been "standardized" by subject, grade and year/semester, meaning that the mean score for each subject in each grade, and year/semester is set to zero and the standard deviation of scores equals one.

5. The formative assessment is the i-Ready exam, which uses a vertical scale from 100–800 that allows comparison of growth within and across years. Because scale-score growth can vary between grade levels, we include grade fixed effects in all our models. Thus, the change in scale scores per instructional day for a student are compared to the achievement growth of other students within the same grade level.

6. For detailed information on disciplinary codes and frequency of each disciplinary incidents by student in the study sample, refer to Appendix Table A2 and Table A3.

7. Individual behavior can vary along many dimensions. However, we take the behavior of peers as given and seek to determine how the impact of unruly peers varies by learning mode.

8. 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.

9. A student was given either a "yellow" flag or a "red" flag, indicating the student took less than 21 or 12 seconds on average, respectively, to finish each task on the exam.

10. The attendance data span from August 17, 2020, through January 25, 2021.

11. For students engaged in full-time in-person learning, attendance was measured in the traditional way. For virtual learners, there was some variation in attendance criteria over time. During the universal remote period at the beginning of the fall 2020 semester and the subsequent transitional hybrid-learning period, an elementary student assigned to remote learning on a given day had to be present during reading or math instruction in order to be counted as present that day. For middle and high school, students assigned to remote instruction had to be present for 50% or more of the school day to be counted as present. From the time that full-time in-person learning was offered in mid-October until the end of the fall semester, elementary students who remained in virtual learning were considered present if they checked in via the Microsoft Teams or i-Ready applications at any point during the day (attendance procedures did not change for secondary students).

12. The relationship between educational inputs and student achievement is formalized in the Appendix, using a cumulative achievement function model. The specification of the empirical model that we estimate is also discussed in detail in the Appendix.

13. As described in the data section, the sample for these summary statistics includes students that (a) took the formative assessments in at least one pre-pandemic academic years (SY 2018–19 or SY 2019–20) and (b) have records during the transitional period (between fall and winter tests in SY 2020–21). Therefore, those who were in Grade 8 during the pre-pandemic academic years and Grade 1 in SY 2020–21 are excluded. Summary statistics of student demographics are presented in Appendix Table A4.

14. Descriptive statistics for the 2SLS estimation sample are in Table A5 of the Appendix.

15. Our approach follows that of Goldhaber et al. (2022).

16. Decomposition results for the 2SLS model appear in Appendix Table A6.

About the Authors

Sungmee Kim

Sungmee Kim was a graduate research assistant with Georgia Policy Labs. She is now a postdoctoral associate at the Samuel DuBois Cook Center on Social Equity at Duke University. She received her Ph.D in economics from Georgia State University. She holds a bachelor's degree in economics from Georgia College & State University and a bachelor's degree in international trade and logistics and a bachelor's degree in international business from Pukyong National University in South Korea. Her research interests are in education and health policy for K–12 students.

Tim R. Sass

Tim R. Sass in a Distinguished University Professor in the Department of Economics at Georgia State University and the W.J. Usery Chair of the American Workplace in the Andrew Young School of Policy Studies. He is also the faculty director of the Metro Atlanta Policy Lab for Education (MAPLE). His research interests include the teacher labor supply, the measurement of teacher quality, and school choice. His work has been published in numerous academic journals and has been supported by several federal and philanthropic grants. He has acted as a consultant to school systems across the country. He is also a senior researcher at the Center for Analysis of Longitudinal Data in Education Research (CALDER).

About the Georgia Policy Labs

The Georgia Policy Labs is an interdisciplinary research center that drives policy and programmatic decisions that lift children, students, and families—especially those experiencing vulnerabilities. We produce evidence and actionable insights to realize the safety, capability, and economic security of *every* child, young adult, and family in Georgia by leveraging the power of data. We work alongside our school district and state agency partners to magnify their research capabilities and focus on their greatest areas of need. Our work reveals how policies and programs can be modified so that every child, student, and family can thrive.

Housed in the Andrew Young School of Policy Studies at Georgia State University, we have three components: the Metro Atlanta Policy Lab for Education (metro-Atlanta K–12 public education), the Child & Family Policy Lab (supporting children, families, and students through a cross-agency approach), and the Career & Technical Education Policy Exchange (a multi-state consortium exploring high-school based career and technical education).

Learn more at gpl.gsu.edu.