NATIONAL CENTER for ANALYSIS of LONGITUDINAL DATA in EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of North Carolina at Chapel Hill, University of Texas at Dallas, and University of Washington

The Consequences of Remote and Hybrid Instruction During the Pandemic

Dan Goldhaber
Thomas J. Kane
Andrew McEachin
Emily Morton
Tyler Patterson
Douglas O. Staiger

WORKING PAPER No. 267-0522 • May 2022

The Consequences of Remote and Hybrid Instruction During the Pandemic

Dan Goldhaber
American Institutes for Research/CALDER
University of Washington/CEDR

Thomas J. Kane
Harvard University

Andrew McEachin
NWEA

Emily Morton
NWEA

Tyler Patterson
Harvard University

Douglas O. Staiger
Dartmouth College

Contents

Acknowledgments

The research was supported by grants from the Carnegie Corporation of New York, The Walton Family Foundation, Kenneth C. Griffin, and CALDER, which is funded by a consortium of foundations (for more information about CALDER funders, see www.caldercenter.org/about-calder).

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any opinions, findings, and conclusions expressed in these papers are those of the authors and do not necessarily reflect the views of our funders. All errors are our own.

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

The Consequences of Remote and Hybrid Instruction During the Pandemic

Dan Goldhaber, Thomas J. Kane, Andrew McEachin, Emily Morton, Tyler Patterson, and Douglas O. Staiger

CALDER Working Paper No. 267-0522
May 2022

Abstract

Using testing data from 2.1 million students in 10,000 schools in 49 states (plus D.C.), we investigate the role of remote and hybrid instruction in widening gaps in achievement by race and school poverty. We find that remote instruction was a primary driver of widening achievement gaps. Math gaps did not widen in areas that remained in-person (although there was some widening in reading gaps in those areas). We estimate that high-poverty districts that went remote in 2020-21 will need to spend nearly all of their federal aid on academic recovery to help students recover from pandemic-related achievement losses.

Introduction

Since the pandemic started in March 2020, multiple reports have highlighted large declines in students' math and reading achievement as well as widening gaps by race and school poverty.[1] If allowed to become permanent, such losses will have major impacts on future earnings and intergenerational mobility.[2] Although the federal government has provided $190 billion in aid to education agencies, the final package of aid was committed in Spring of 2021 before the impact of the pandemic on achievement was clear. The American Rescue Plan only required districts to spend 20 percent on academic recovery.

We use student-level data from 2.1 million students in 10,000 schools from 49 states (plus D.C.) to compare students' achievement growth during the pandemic (Fall 2019 to Fall 2021) to a pre-pandemic period (Fall 2017 to Fall 2019). In addition to documenting the magnitude of the learning loss, we investigate the role of remote and hybrid instruction in widening gaps in achievement by race and school poverty. A prior study by Jack et al. (2021) documented declines in proficiency rates in districts that shifted to remote instruction, especially in districts serving larger shares of Black and Hispanic students and lower income students. However, without access to within-district comparisons, their work could not distinguish between a true differential impact on disadvantaged students and district-wide differences for districts serving larger shares of low-income students (e.g. in the implementation of remote instruction.) Their study was also limited to 12 states.[3]

We make five primary contributions: First, we estimate a model of achievement growth in the pre-pandemic period (conditioning on student and school characteristics as well as prior achievement) and then compare students' actual and expected achievement growth during the

[1] For instance, see: Curriculum Associates (2020, 2021a, 2021b); Darling-Aduana et al. (2021); Dorn et al. (2020); Kogan and Lavertu (2021); Kuhfeld et al. (2021); Lewis and Kuhfeld (2021); Lewis et al. (2021).
[2] Using evidence on test scores and achievement from Neal and Johnson (1996) and Murnane, Willett and Levy (1995), Goldhaber, Kane and McEachin (2021) estimated that the losses would cost the U.S. $2 trillion in lifetime earnings. The World Bank estimated that the worldwide losses in lifetime earnings would be $17 trillion (Azevedo et al., 2022).
[3] Their primary outcome is proficiency on state tests. Because states have different proficiency standards and different shares of students near those standards, their study could indicate the direction but not the magnitude of the impact. Kilbride et al. (2021) also find larger declines in achievement in schools that went remote in the state of Michigan.

pandemic. By doing so, we distinguish pandemic-related achievement losses from pre-existing differences in achievement growth by student and school characteristics.

Second, we investigate differential impacts on high and low-income schools when their districts shifted to remote instruction. We find that the shift in instructional mode was a primary driver of widening achievement gaps by race/ethnicity and by school poverty status. Within school districts that were remote for most of 2020-21, high-poverty schools experienced 50 percent more achievement loss than low-poverty schools (e.g. .46 vs. .30 standard deviations in math.) In contrast, math achievement gaps did not widen in areas that remained in-person (although there was some widening in reading gaps in those areas).

Third, after documenting higher rates of remote instruction in high poverty schools, we decompose the role played by the differing incidence and differing impacts of remote instruction. High poverty schools were more likely to go remote and they suffered larger declines when they did so. Although the former played a role, the latter was more important.

Fourth, we investigate within-school differences in the impacts of the pandemic on student subgroups. We find that most of the widening by race/ethnicity occurred because the schools attended by Black and Hispanic students were more negatively impacted, rather than because they fell behind classmates attending the same school. Put another way: the widening racial gap happened because of negative shocks to schools attended by disadvantaged students, not because of differential impacts within schools.

Fifth, we provide a lower bound estimate of the cost of academic recovery by district. To do so, we compare the share of a typical school year that students have lost to the share of their annual budget they have received in federal aid. Such an estimate is likely to be a lower bound, as long as the marginal cost per unit of achievement growth is higher for catch-up efforts than during the typical school year. We estimate that high poverty districts that were remote for most of 2020-21 will need to spend nearly all of their federal aid on academic recovery in order to eliminate the losses their students have experienced.

Student Achievement Data

For a national sample of student achievement, we rely on data from the Growth Research Database (GRD) of NWEA, a non-profit assessment provider. Roughly three thousand school districts administer NWEA's Measures of Academic Progress (MAP) Growth assessments. Unlike state-mandated tests, districts typically administer the MAP assessment three times per year: in the Fall, Winter, and Spring. Though some remote testing occurred during the pandemic, nearly all MAP Growth tests were administered in-person at the students' schools during the three fall terms included in the present study.

The MAP Growth assessment is a computer adaptive test, meaning that the difficulty of test questions increases or decreases in response to a student's prior responses. In contrast to tests with a standard test form for all students, the adaptive tests are designed to improve reliability at both the high end and low end of achievement. Test scores are computed based on the Rasch item response theory (IRT) model, and the tests are vertically scaled so that scores can be meaningfully compared across different grades.

The NWEA test is ideal for measuring achievement during the pandemic, since so many students are scoring below their current grade level. We have standardized scores using the means and standard deviations by grade and subject using NWEA's most recent pre-pandemic norms[4] (Thum and Kuhfeld, 2020). In our analyses, we also control for testing dates. The NWEA data also include student-level demographic data on race/ethnicity and gender, as well as district and school identifiers.

We supplement the NWEA data with administrative data from the Common Core of Data (CCD): enrollment by school and grade in 2019-20, the urbanicity of the school, expenditures on elementary and secondary education, and the percent of students in each school qualifying for the federal Free- and Reduced-Price Lunch Program.[5] In addition to the CCD, we added

[4] The NWEA national norms have been weighted to reflect the national population of K-8 public schools in 2015-16. The means and standard deviations were estimated pooling data over three school years, 2015-16, 2016-17 and 2017-18.

[5] Where FRPL values were unavailable, we used the percent of students meeting eligibility for federal lunches through direct certification. This included the entirety of three states (DE, MA, and DC), as well as 2.6 percent of schools outside these states. We also added data from the American Community Survey on the characteristics of the

information on the population density (population per square mile) within each school district using data from the Census Bureau, COVID infection rates by county from Johns Hopkins University[6] and estimates of federal Elementary and Secondary School Emergency Relief (ESSER) Funds by district.[7]

To measure schools' instructional mode during 2020-21, we rely on the *Return to Learn Tracker* assembled by the American Enterprise Institute (AEI).[8] The AEI data include weekly data on mode of instruction from August 2020 through June 7, 2021 for 98 percent of enrollment in U.S. school districts with 3 or more schools.[9]

Representativeness of the Analysis Sample

Our analytic sample for math consists of 2.1 million students at 9,692 schools from 49 states (plus D.C.)[10]. The sample includes students who were in grades 3 to 8 in the follow-up year. We included schools that were covered in the AEI data and had valid test scores for at least 10 students on the English language versions of the mathematics or reading assessments in Fall 2017, Fall 2019, and Fall 2021 (all three years). In addition, individual students were required to have scores for both a baseline year (i.e., Fall 2017 or Fall 2019) and a follow-up test two years

population within school boundaries using the School Attendance Boundary Survey of 2015-16, such as the percent of households with broadband access, adult employment in wholesale and retail trade and health professions. None of the results are sensitive to including them as covariates.

[6] The Covid infection rate data is compiled by Johns Hopkins' Center for Systems Science and Engineering and is available at https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series.

[7] We estimated ESSER allocations by district using state ESSER totals and prior Title 1 allocations for each district. The federal legislation required states to allocate 90 percent of the ESSER funds using Title 1 spending in FY2019 and FY2020.

[8] Given missing data in the early weeks, we start from September 7, 2020, the date for which over 95% of available districts have data.

[9] To identify the effects of instructional mode, we needed to know the school a student attended during the academic year preceding the Fall follow-up assessment. Most students participated in at least one assessment during the intervening year (2018-19 or 2020-21) and we used the testing data to link to schools. If students attended the same school in the baseline and follow-up year, we assume they attended that school during the intervening year. For students who changed schools between the baseline and follow-up year (and advanced two grade levels), we use grade-span data for their former and current school. For example, we assume that a fourth grader at a K-5 elementary school in Fall 2019 who was a 6th grader at a 6-8 middle school in Fall of 2021 would have been a 5th grader in the elementary school. In instances in which both schools serve the student's grade level in the intervening year, we treat the school as missing.

[10] The NWEA analysis file only included scores for students taking the English language version of the test.

later (i.e., Fall 2019 and Fall 2021). Finally, students were excluded if their school tested less than sixty percent of their grade's enrollment based on data from the CCD.

In Appendix Table 3, we report descriptive statistics for our analysis samples as well as for the full CCD universe of public schools with students in grades 3-8. In comparison to the national population, our analytic sample for studying math achievement contains a smaller percentage of Hispanic students (20 percent vs. 28 percent nationally), slightly less representation of high poverty schools (22 percent versus 27 percent) and greater representation among suburban schools (44 percent versus 39 percent) than the national population of public schools. The analytic sample also had similar percentages of the year spent in remote and hybrid instruction (21 and 47 percent respectively) as for all schools with both CCD and AEI data (24 and 46 percent).

The requirement that students have a follow-up score led us to exclude roughly a quarter of students with valid baseline tests (25 percent in Fall 2017 and 29 percent in Fall 2019).[11] In Appendix Table 5, we report the degree to which each of the covariates is related to attrition in both the pre-pandemic and pandemic periods. Given the change in attrition rates, we test the robustness of our findings by including the share of students tested in the school as a covariate.

Differing Incidence of Remote Instruction by School Poverty Level

As others have found (Parolin and Lee, 2021; Camp and Zamarro, 2021; Grossmann et al., 2021; Oster et al., 2021), we observe a higher incidence of remote schooling for Black and Hispanic students. We also find that high poverty schools spent about 5.5 more weeks in remote instruction during 2020-21 than low and mid poverty schools.[12]

We observed large differences in remote instruction in different states. In Figure 1, we sort states into four categories based on percentage of students in remote instruction. High

[11] Further excluding the students at schools whose schools tested less than 60 percent of their grade's enrollment dropped 0.3 and 2.2 percent of students in the NWEA and AEI sample who respectively tested in follow-up years Fall 2019 and Fall 2021.
[12] We investigated whether the higher incidence of remote instruction in high-poverty schools was due to greater population density, the urbanicity of the school (which varies especially within countywide school districts) and higher COVID infection rates in the county. After adjusting for such factors, the gap in weeks of remote instruction between high and low-poverty schools is only slightly smaller (roughly 4.6 weeks).

poverty schools were more likely to be remote in all four groups of states, but the gaps were largest in those states with higher rates of remote instruction overall. For example, in high remote instruction states (including populous states such as California, Illinois, New Jersey, Virginia, Washington and the District of Columbia), high-poverty schools spent an additional 9 weeks in remote instruction (more than 2 months) than low-poverty schools. In states with the lowest rates of remote instruction (including populous states such as Florida and Texas), high poverty schools were again more likely to be remote, but the differences were small: 3 weeks remote in high poverty schools versus 1 week remote in low poverty schools.[13]

Inferring the Impacts of Remote and Hybrid Instruction

As noted in the introduction, we compare student achievement growth during the pandemic (Fall 2019 to 2021) to growth expectations from a pre-pandemic period (Fall 2017 to 2019). To establish pre-pandemic growth expectations, we first estimate the following model of achievement growth (Todd and Wolpin, 2003) during the pre-pandemic period:

$$\mathrm{S_{i0}} = \beta_0 + \mathrm{Race_i}\beta_{\mathrm{Race}} + \mathrm{Pov}_{j0}\beta_1 + Mode_{j,2021}\beta_2 + \mathrm{Pov}_{jo}Mode_{j,2021}\beta_3 + X_{ij0}\beta_4 + \varepsilon_{i0}$$

where i subscripts the student, j subscripts the school attended in 2018-19 (the school year between the baseline year and follow-up) and the zero subscript refers to the pre-pandemic period. $\mathrm{Race_i}$ is a vector of dummies for students' race/ethnicity (Black, Hispanic, Asian, and Other with White as the reference group), Pov_{j0} is a vector of dummies for the poverty status of the school attended (mid and high poverty with low poverty as the reference group), $Mode_{j,2021}$ is a vector with the percentage of the year that a school was hybrid and remote during the 2020-21 school year, and $X_{ij,0}$ is a vector of student and school characteristics (including a cubic in baseline achievement fully interacted with grade level, gender and the date of testing in the baseline and in the follow-up year included as linear terms).

The parameter estimates (reported in Appendix Table 4) reveal that even before the pandemic, there were significant differences in achievement growth by race/ethnicity and school poverty status after controlling for baseline achievement. For example, relative to white students

---

[13] States with low closure rates included Arkansas, Florida, Idaho, Louisiana, Maine, Montana, North Dakota, Nebraska, South Dakota, Texas, Utah, Vermont and Wyoming.

with similar baseline scores and school poverty levels, Black students' math test scores were .12 standard deviations lower two years later, and Hispanic students' scores were .02 standard deviations lower. The magnitude of widening for Black and Hispanic students was similar in reading. Conditioning on student race/ethnicity and baseline scores, students in high poverty schools also fell behind by approximately .18 standard deviations in math and .14 standard deviations in reading during 2017-19.

In the growth model above, we also included controls for the instructional mode used by their intervening year's school during the 2020-21 school year. Although there should be no causal relationship between remote/hybrid schooling in 2020-21 and student growth between 2017-19, we estimate such differences to identify any pre-existing relationships between a school's subsequent use of remote/hybrid schooling and growth. The differences were small but, in some cases, statistically significant. As described below, we difference those out from 2019-21 growth.

Thus, model (1) above establishes a benchmark for how achievement, conditional on prior scores, varied by race, school poverty and pandemic instructional mode before the pandemic. We use those estimates to construct our primary outcome, which is the degree to which each student in 2019-21 underperformed (or overperformed) growth expectations from the 2017-19 period.[14] Specifically, we apply the 2017-19 parameters to the 2019-21 sample to estimate the difference between a student's actual and expected growth during the pandemic as follows:

$$(1)\ \mathrm{R}_{i1} = \mathrm{S}_{\mathrm{i1}} - (\hat{\beta}_0 + \mathrm{Race}_{\mathrm{i1}}\hat{\beta}_{\mathrm{Race}} + \mathrm{Pov}_{j1}\hat{\beta}_1 + Mode_{i,j2021}\hat{\beta}_2 + \mathrm{Pov}_{j1}Mode_{j,2021}\hat{\beta}_3 + X_{ij1}\hat{\beta}_4)$$

Thus, when we refer to a "loss" or "decline" in achievement growth, we mean that actual achievement growth was less than expected given pre-pandemic relationships ($\mathrm{R}_{i1} < 0$).

[14] Our approach to measuring growth is different from that used by NWEA in its national reports. In estimating growth norms, NWEA conditions on baseline scores, testing date and grade—but not race/ethnicity or school poverty level. Thus, since there were pre-existing differences in achievement growth by race/ethnicity or school poverty, they are included in the pandemic learning losses for such groups.

In the discussion below, we will focus on math achievement while providing analogous analyses for reading achievement in an appendix.[15] Although magnitudes are smaller, the pattern of results are similar in reading—with one important exception which we highlight when discussing Tables 1 and 2 below. For brevity, we also pool results across grades 3 through 8. Although the magnitudes of differences are larger in grades 3-5 than in 6-8, the patterns are similar.[16]

In Table 1, we describe how 2020-21 growth diverged from expectations for different subgroups of students by regressing R_{i1} on different combinations of covariates.[17] In column 1, we report that Black and Hispanic students lost even more ground relative to White students with similar baseline achievement during the pandemic period than in the pre-pandemic period: Black students lost an additional .119 standard deviations and Hispanic students lost an additional .092 standard deviations. (As reflected in the constant term, White students, the excluded subgroup, also lost .208 standard deviations relative to the pre-pandemic period.)

In column (2), we report differences in R_{i1} by students' baseline achievement. As reflected in the constant term, actual growth for students in the highest quartile on the baseline assessment (the excluded category) during the pandemic period was .194 standard deviations lower than expected growth. Students who were in the middle two quartiles of achievement in Fall 2019 lost an additional .053 standard deviations, while students in the bottom quartile in the baseline lost an additional .107 standard deviations.

In column (3), we report the gaps by race and by baseline score while conditioning on both student characteristics. Because student race/ethnicity and baseline score are correlated, the magnitude of the loss for each is somewhat smaller when conditioning on both.

In column (4), we include school fixed effects. Although they are still positive, the Black-White and Hispanic-White achievement gaps in math achievement are greatly diminished by the

[15] Appendix Tables 1 and 2 contain reading analogues to Tables 1 and 2, respectively. Appendix Figures 1 and 2 contain reading analogues to Figures 2 and 3, respectively. Other appendix tables contain math and reading results side-by-side.
[16] In math, the pattern of results by race, school poverty and instructional mode are similar in elementary and middle school grades. In reading, the direct effect of school poverty status is larger in middle school grades.
[17] In Appendix Table 1 we present analogous results for reading.

inclusion of school fixed effects, falling to .036 and .032 standard deviations respectively. The smaller magnitudes suggest that much of the increased gap in test scores reported in column (3) is a result of school-level shocks rather than differential effects of the pandemic on racial/ethnic subgroups within schools. Likewise, the gap in math achievement between students in the highest and lowest quartile of baseline achievement shrinks by 72 percent with the inclusion of school fixed effects (.022/.078=.28).

The results in column (4) have implications for academic recovery efforts: to reverse pandemic-related losses (as opposed to addressing long-standing inequities) districts might focus on the hardest hit schools, rather than target subgroups within schools.

In column (5), we parameterize school effects on math achievement with three factors: the school poverty status (low-poverty, mid-poverty and high-poverty), the percentage of the 2020-21 school year that the school was in remote or hybrid instruction, and the interaction between school poverty status and instructional mode.[18] The conditional difference by race/ethnicity remains small, implying that the simple parameterization captures much of the information in the school effects specification in column (4).[19]

Several other findings from Table 1 are noteworthy. In column (5), the main effects of school poverty status—which apply to those schools that were in-person for all of 2020-21—are small and no longer statistically significant. In other words, as long as schools were in-person throughout 2020-21, there was no widening of math achievement gaps between high-, middle-, and low-poverty schools.

The main effects of hybrid and remote instruction are negative, implying that even at low-poverty (high income) schools, students fell behind growth expectations when their schools went remote or hybrid. Specifically, if their schools were remote throughout 2020-21, students

[18] Investigating further, we found that the variance in school effects increased by 81 percent between 2017-19 and 2019-21, as schools were differentially impacted during the pandemic. However, when we controlled for three variables (school poverty status, the percent of weeks remote/hybrid and the interaction,) the variance in school effects largely returned to levels seen in 2017-19. That is, the parameterization seemed to account for between 57 and 66 percent of the increase in variance (See Appendix A.)

[19] The differences by baseline score bounce back partially between columns (4) and (5), but remain far smaller than those in column (3). Apparently, the schools attended by low-baseline score students are different in ways not captured by school poverty status or by percent remote/hybrid.

in low-poverty schools lost .201 standard deviations relative to expected growth. The losses associated with hybrid instruction were smaller, equal to .033 standard deviations if schools were hybrid the whole year.

Perhaps the most striking finding in column (5) is that the consequences of hybrid and remote instruction for math achievement were substantially larger in mid- and high-poverty schools than in low-poverty schools: the interaction between percent remote and high poverty was -.158, which means high poverty schools that were remote all year lost .359 standard deviations (-.158-.201) more than high poverty schools that were in person all year. High poverty schools spending the year in hybrid instruction lost .150 standard deviations (-.033-.117) relative to high poverty schools that remained in person. When we focus on within-district differences (by including district fixed effects in column (6)), the losses associated with remote and hybrid instruction remained similar for mid and high poverty schools.

In column (7), we adjust for attrition by including the ratio of the number of tested students in the school to the number of students enrolled in the relevant grades in the school during the 2019-20 school year. The substantive results are unchanged.

In Figure 2, we report the mean of R_{i1} by the percentage of the year schools were in remote instruction and by school poverty (conditioning on the covariates in Table 1). The vertical axis intercepts for the three lines are similar, implying that among those schools that were not remote during 2020-21, the losses were similar for low, medium and high-poverty schools—about .17 standard deviations on average. Presumably, such losses reflect some combination of the disruptions during Spring 2020 (when all schools spent time in remote instruction) and the effect of pandemic-related stresses during 2020-21. However, the gaps between high and low poverty schools are wider for schools that spent a larger share of the year in remote and hybrid instruction. For schools that spent more than 50 percent of the year in remote instruction, students in high poverty schools lost roughly .44 standard deviations relative to pre-pandemic growth, while students in low-poverty schools lost .26 standard deviations.

In Appendix Table 1, we report similar findings for students' reading scores. In terms of standard deviation units, the losses were smaller, but we see the same pattern of small racial/ethnic losses within schools and larger impacts of remote and hybrid schooling on students attending mid and low-poverty schools. However, one substantive difference between math and reading is that gaps in reading achievement by school poverty and race did widen somewhat in districts which remained in person. While students learn math primarily in school, student learning in reading may depend more on parental engagement at home. Thus, the contrast between the math and reading findings for in-person districts may reflect differential family stresses outside of school.

Disparate Incidence vs. Disparate Impact of Remote and Hybrid Schooling

High poverty schools were more likely to go remote and the consequences for student achievement were more negative when they did so. Which was more important? In Table 2, we decompose the role played by the two factors-- disparate incidence and disparate impacts-- in widening the gap between low and high poverty schools.[20] In the top row, we report the total difference in actual vs. expected math achievement gains between high and low-poverty schools, which is .168 standard deviations. As reported in the next two rows, a small share of this difference (.014+.016) was due to the fact that Black and Hispanic students and students with low baseline achievement scores gained less, and that those students were more likely to attend high poverty schools. In the fourth row, we add in the differential loss in achievement gains between high and low poverty schools in areas that were in-person throughout 2020-21. As noted earlier, there was essentially no widening in math achievement gaps in districts that were fully in-person (.002 standard deviations). In the fifth row, we report the effect of greater incidence of remote/hybrid instruction in high-poverty schools, which was about one third of the total difference (.051/.168). The remaining half of the gap (.085/.168) was due to the differing impact of hybrid/remote instruction on high poverty schools. (We describe the methodology for decomposition in Appendix B.)

As reported in Appendix Table 2, a larger share of the widening gap in reading achievement between high and low-poverty schools was due to widening gaps in areas that remained in person (26 percent). Accordingly, the shares that were due to disparate incidence

[20] We describe the algebra for the decomposition in Appendix B.

(19 percent vs. 30 percent) and disparate impacts of remote/hybrid instruction (35 percent vs. 51 percent) were lower in reading than in math.

Paying for Academic Recovery

From the beginning of the pandemic through to the American Rescue Plan in Spring 2021, the federal government provided state and local education agencies with $190 billion to pay for COVID-related expenses. States are required to allocate 90 percent of that funding to districts based on the Title I formula, which reflects child poverty rates and public assistance receipt in each district. Importantly, the funds were committed before the impact of the pandemic and instructional mode were clear. In this section, we provide a simple rule of thumb for judging whether the federal dollars are likely to be sufficient to pay for the catch-up in each district.

To put the achievement impacts and the federal aid on a comparable scale, we convert each into the share of each district's annual budget they represent. It is straightforward to convert the federal aid into an annual budget share, dividing each district's allocation by its spending on K-12 education in 2019-20 (minus capital expenditures.)

To convert recovery costs into an annual budget share, we estimate the share of a typical school year (in terms of instructional weeks) that would be required to make up for lost achievement during the pandemic. The NWEA data are especially well-suited to this task. Unlike the official state tests, school districts implement the the NWEA's MAP assessment at different points on their academic calendars. Thus, the test developers have observed how scores vary by the number of instructional weeks students received between test dates (which would yield unbiased estimates of gains per week of instruction as long as timing is exogenous; Thum and Kuhfeld, 2020).[21] After using the parameters in column (5) of Table 1 to estimate each school's reduction in math test score gains, we divide by an estimate of instructional growth in math per week for grades 3 through 8 from NWEA to estimate the number of instructional weeks required for schools to get back to pre-pandemic growth expectations. To translate the estimated

[21] Because the tests are given in the Spring and in the Fall, the gains per instructional week during the school year do not include summer learning loss.

weeks into a portion of the school year, we then divide the estimate of lost weeks by 40 (the number of calendar weeks in the typical school year) and aggregate to the district level (where ARP spending decisions will be made).[22]

The share of a district's annual budget equivalent to the share of a typical school year missed is likely to be a conservative estimate of the cost of recovery.[23] To make up 20 percent of a school year's worth of unfinished learning, it is likely to cost *more than* the equivalent of 20 percent of a district's annual budget. For instance, imagine if a district extends the school year or lengthens the school day. They are likely to have to pay teachers more than their normal wage rate (e.g., "time and a half") and, if students or teachers are tired at the end of the day or year, the marginal learning gain from additional time is likely to smaller as well. While many schools are exploring alternative ways of organizing instruction—e.g., with small group tutoring—the marginal cost per a given gain in achievement for these alternative models is likely to be more than under the predominant technology of schooling (e.g., with 20-25 students per elementary teacher).

The correlation between the share of a year of unfinished learning and the share of an annual budget received in federal aid is positive (.35), largely because both are positively related to poverty.

In Figure 3, we compare the shares of a school year required to eliminate the achievement loss and shares of annual budgets represented by federal aid. We do so for four categories of schools. On the left are school districts that have below-median percentages of students receiving federal free lunches; on the right are the above-median (higher poverty) districts. Within each, we report separately for districts that were fully in-person during 2020-21 and for

[22] We assume that district operational expenditures are spread over 40 calendar weeks, rather than the 36 instructional weeks (180 days) that is the norm in most states. If we were to use instructional weeks, the estimated cost of recovery would by roughly 10 percent larger. Providing instruction outside the traditional classroom format of 20 to 25 students per teacher in an elementary school—e.g. tutoring or after-school programs—is likely to cost more per s.d. of achievement gain. Otherwise, it would be difficult to explain the ubiquity of the traditional classroom model.

[23] An alternative approach would be to start with various types of interventions—such as tutoring and after school and extra periods of math instruction—for which we have credible impact estimates and estimate what it would cost to eliminate the gaps observed. However, one would have to make additional assumptions about the cost and efficacy of a dramatic scale-up of those programs. Tutor salaries are likely to vary by local labor market conditions.

those that spent the majority of the year remote or hybrid. (For brevity, we excluded districts between the two extremes, who were remote/hybrid for less than half the year.)

Ironically, it is the lower-poverty districts choosing to remain remote during 2020-21 who face the greatest shortfall. Because the federal aid was based on the Title I formula, the lowest poverty (highest income) public school districts received less than 15 percent of their annual budgets in federal aid. The low-poverty districts who were remote or hybrid for most of the year lost 27 percent of a year's learning.

On the right, we compare federal aid and academic losses for the highest poverty quartile districts (lowest incomes). For high-poverty districts that remained in person, the losses were similar to those of low-poverty schools that remained in person (about 15 percent of a school year). However, because the federal dollars were based on poverty and not their achievement losses, they received considerably *more* funding (about a third of their annual budgets) than the 15 percent of a school year of unfinished learning their students experienced.

On the far right, we report the average losses for high-poverty districts that remained remote. The hardest hit group, their lost achievement amounted to slightly under 40 percent of a year of learning. That is roughly equivalent to the share of their annual budgets they received in federal aid.

The American Rescue Plan only *requires* districts to spend 20 percent on academic recovery. According to an analysis of district plans by the non-profit, Future-Ed, at Georgetown University, the average district is planning to spend not much more than the minimum on academic recovery (28 percent), with the remainder planned for facilities, technology, staffing and mental and physical health.[24]

Conclusion

Throughout the country, local leaders made different choices about whether to hold classes in-person or remotely during the COVID-19 pandemic. There were valid reasons for differing judgements—including differing risks related to local demographics or population

[24] https://www.future-ed.org/financial-trends-in-local-schools-covid-aid-spending/

density as well as real uncertainty about the public health consequences of in-person schooling. While we have nothing to add regarding the public health benefits, it seems that the shifts to remote or hybrid instruction during 2020-21 had profound consequences for student achievement. In districts that went remote, achievement growth was lower for all subgroups, but especially for students attending high-poverty schools. In areas that remained in person, there were still modest losses in achievement, but there was no widening of gaps between high and low-poverty schools in math (and less widening in reading.)

It is possible that the relationships we have observed are not entirely causal, that family stress in the districts that remained remote both caused the decline in achievement and drove school officials to keep school buildings closed. However, even if that were the case, our results highlighting the differential losses in high poverty schools that went remote are still critical for targeting recovery efforts.

While local leaders are well aware of the losses in student achievement, they have received little guidance when translating declines in math and reading achievement (typically measured in proficiency rates or percentile points) into an implied scale of recovery effort. We propose one relevant benchmark—the share of a typical school year that would be required to make up for the loss. It is a lower bound estimate, since the marginal cost per unit of growth from supplemental recovery efforts is likely to be higher than the average cost during a typical school year. Another approach is to convert the achievement loss into standard deviation units to facilitate comparison with the effect sizes for relevant interventions. For instance, the average high poverty school that remained in remote instruction for a majority of 2020-21 lost roughly .44 standard deviations in achievement. For comparison, a recent review of pre-pandemic research by Nickow et al. (2020) on high-dosage tutoring—defined as tutors working with fewer than 4 students, 3 to 5 times per week for at least 30 minutes—produced a .38 standard deviation gain in math. Thus, in high poverty schools that remained remote, leaders could provide high-dosage tutoring to *every* student still not make up for the loss.

Depending on whether they remained remote during 2020-21, some school agencies have much more work to do now than others. If the achievement losses become permanent, there will be major implications for future earnings, racial equity and income inequality, especially in states where remote instruction was common.

References

Azevedo,Joao Pedro Wagner De; Rogers,F. Halsey; Ahlgren,Sanna Ellinore; Cloutier,Marie-Helene; Chakroun,Borhene; Chang,Gwang-Chol; Mizunoya,Suguru; Reuge,Nicolas Jean; Brossard,Matt; Bergmann,Jessica Lynn. 2022. *The State of the Global Education Crisis : A Path to Recovery*. Washington, D.C. : World Bank Group.

Camp, Andrew, and Gema Zamarro. 2021. “Determinants of Ethnic Differences in School Modality Choices During the COVID-19 Crisis.” *Educational Researcher* 51, no. 1: 6–16.

Curriculum Associates. 2020. “Understanding Student Needs: Early Results from Fall Assessments.” Research brief, Curriculum Associates.

Curriculum Associates. 2021a. “What We’ve Learned about Unfinished Learning: Insights from Midyear Diagnostic Assessments.” Research brief, Curriculum Associates.

Curriculum Associates. 2021b. “Academic Achievement at the End of the 2020–2021 School Year: Insights after More Than a Year of Disrupted Teaching and Learning.” Research brief, Curriculum Associates.

Darling-Aduana, Jennifer, Henry T. Woodward, Tim R. Sass, and Sarah S. Barry. 2021. “Learning-Mode Choice, Student Engagement, and Achievement Growth During the COVID-19 Pandemic.” CALDER Working Paper No. 260-0122, January 2021. https://caldercenter.org/sites/default/files/CALDER%20WP%20260-0122.pdf.

Dorn, Emma, Brian Hancock, Jimmy Sarakatsannis, and Ellen Viruleg. 2020. “COVID-19 and Learning Loss—Disparities Grow and Students Need Help.” McKinsey & Company, December 8, 2020. https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help.

Goldhaber, Dan, Thomas Kane, and Andrew McEachin. 2021. “Analysis: Pandemic Learning Loss Could Cost U.S. Students $2 Trillion in Lifetime Earnings. What States & Schools Can Do to Avert This Crisis.” *The 74 Million*, December 13, 2021. https://www.the74million.org/article/analysis-pandemic-learning-loss-could-cost-u-s-students-2-trillion-in-lifetime-earnings-what-states-schools-can-do-to-avert-this-crisis/.

Grossmann, Matt, Sarah Reckhow, Katharine O. Strunk, and Meg Turner. 2021. “All States Close but Red Districts Reopen: The Politics of In-Person Schooling During the COVID-19 Pandemic.” *Educational Researcher* 50, No. 9: 637–648.

Jack, Rebecca, Clare Halloran, James Okun, and Emily Oster. 2021. “Pandemic Schooling Mode and Student Test Scores: Evidence from U.S. States.” NBER Working Paper No. 29497, December.

Kilbride, Tara, Bryant Hopkins, Katharine O. Strunk, and Scott Imberman. 2021. *K-8 Student Achievement and Achievement Gaps on Michigan’s 2020-21 Benchmark and Summative Assessments*, Education Policy Innovation Collaborative, East Lansing, MI.

Kogan, Vladimir, and Stéphane Lavertu. 2021. “The COVID-19 Pandemic and Student Achievement on Ohio’s Third-Grade English Language Arts Assessment.” The Ohio State University, January 27, 2021. https://glenn.osu.edu/sites/default/files/2021-09/ODE_ThirdGradeELA_KL_1-27-2021.pdf.

Kuhfeld, Megan, Erik Ruzek, Karyn Lewis, Jim Soland, Angela Johnson, Beth Tarasawa, and Lindsay Dworkin. 2021. “Understanding the Initial Educational Impacts of COVID-19 on Communities of Color.” NWEA research report, June 2021.

Lewis, Karyn, and Megan Kuhfeld. 2021. “Learning During COVID-19: An Update on Student Achievement and Growth at the Start of the 2021-22 School Year.” NWEA research brief, December 2021.

Lewis, Karyn, Megan Kuhfeld, Erik Ruzek, and Andrew McEachin. 2021. “Learning During COVID-19: Reading and Math Achievement in the 2020-21 School Year.” NWEA research brief, July 2021.

Murnane, Richard J., John B. Willett and Frank Levy. 1995. “The Growing Importance of Cognitive Skills in Wage Determination” *The Review of Economics and Statistics* Vol. 77, No. 2 (May, 1995), pp. 251-266.

Neal, Derek A. and William R. Johnson. 1996. “The Role of Premarket Factors in Black-White Wage Differences”, *Journal of Political Economy*, Vol. 104, No. 5 (October), pp. 869-895.

Nickow, A., Oreopoulos, P., & Quan, V. 2020. The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence (Working Paper 27476). Cambridge, MA: National Bureau of Economic Research.

Nomi, T. and E. Allensworth (2009) "Double-dose algebra as an alternative strategy to remediation: Effects on students' academic outcome" *Journal for Research on Educational Effectiveness* Vol. 2, pp. 111–148.

Oster, Emily, Rebecca Jack, Clare Halloran, John Schoof, Diana McLeod, Haisheng Yang, Julie Roche and Dennis Roche. 2021. "Disparities in Learning Mode Access Among K–12 Students During the COVID-19 Pandemic, by Race/Ethnicity, Geography, and Grade Level—United States, September 2020–April 2021." Morbidity and Mortality Weekly Report 70, No. 26: 953-958.

Pane, John F., Daniel F. McCaffrey, Shannah Tharp-Taylor, Gary J. Asmus, and Billy R. Stokes. 2006. "Student displacement in Louisiana After the Hurricanes of 2005: Experiences of Public Schools and their Students." RAND Gulf States Policy Institute Technical Report. Santa Monica, November.

Parolin, Zachary, and Emma K. Lee. 2021. "Large Socio-Economic, Geographic and Demographic Disparities Exist in Exposure to School Closures." Nature Human Behavior 5, No. 4: 522-528.

Taylor, Eric. 2014. "Spending more of the school day in math class: Evidence from a regression discontinuity in middle school," Journal of Public Economics, Volume 117, pp. 162-181.

Thum, Yeow Meng, and Megan Kuhfeld. 2020. "NWEA 2020 MAP Growth Achievement Status and Growth Norms for Students and Schools." NWEA white paper, July 2020.

Todd, Petra and Kenneth Wolpin. 2003. "On the specification and estimation of the production function for cognitive achievement" The Economic Journal 113 (485), F3-F33.

Table 1. Pandemic Achievement Gains by Student and School Characteristics, Math

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Race (Reference: White)							
Black	-0.119		-0.101	-0.036	-0.040	-0.057	-0.040
	(0.012)		(0.011)	(0.004)	(0.007)	(0.005)	(0.007)
Hispanic	-0.092		-0.077	-0.032	-0.014	-0.043	-0.014
	(0.015)		(0.015)	(0.003)	(0.007)	(0.004)	(0.007)
Asian	-0.013		-0.020	-0.029	0.005	-0.026	0.005
	(0.013)		(0.013)	(0.006)	(0.010)	(0.007)	(0.010)
Other	-0.041		-0.035	-0.019	-0.017	-0.025	-0.017
	(0.009)		(0.009)	(0.003)	(0.009)	(0.004)	(0.009)
Baseline Score (Reference: Top Quartile)							
Middle Quartiles		-0.053	-0.040	-0.012	-0.030	-0.016	-0.030
		(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bottom Quartile		-0.107	-0.078	-0.022	-0.053	-0.030	-0.053
		(0.008)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
School Poverty (Reference: Low <25%)							
Middle (25%-75%)					-0.018	0.020	-0.017
					(0.014)	(0.014)	(0.014)
High (>75%)					-0.002	0.024	-0.001
					(0.019)	(0.019)	(0.019)
Remote Schooling							
% Remote in 2020-21					-0.201	N/A	-0.199
					(0.035)		(0.034)
Interactions:							
• Middle Poverty					-0.086	-0.103	-0.086
					(0.034)	(0.023)	(0.034)
• High Poverty					-0.158	-0.183	-0.159
					(0.037)	(0.030)	(0.037)
Hybrid Schooling							
% Hybrid in 2020-21					-0.033	N/A	-0.033
					(0.019)		(0.018)
Interactions:							
• Middle Poverty					-0.051	-0.023	-0.051
					(0.020)	(0.021)	(0.020)
• High Poverty					-0.117	-0.084	-0.119
					(0.032)	(0.029)	(0.033)
% Tested in School							0.027
							(0.033)
Constant	-0.208	-0.194	-0.175	N/A	-0.098	N/A	-0.122
	(0.006)	(0.006)	(0.006)		(0.014)		(0.033)
Fixed Effects?	No	No	No	School	No	District	No

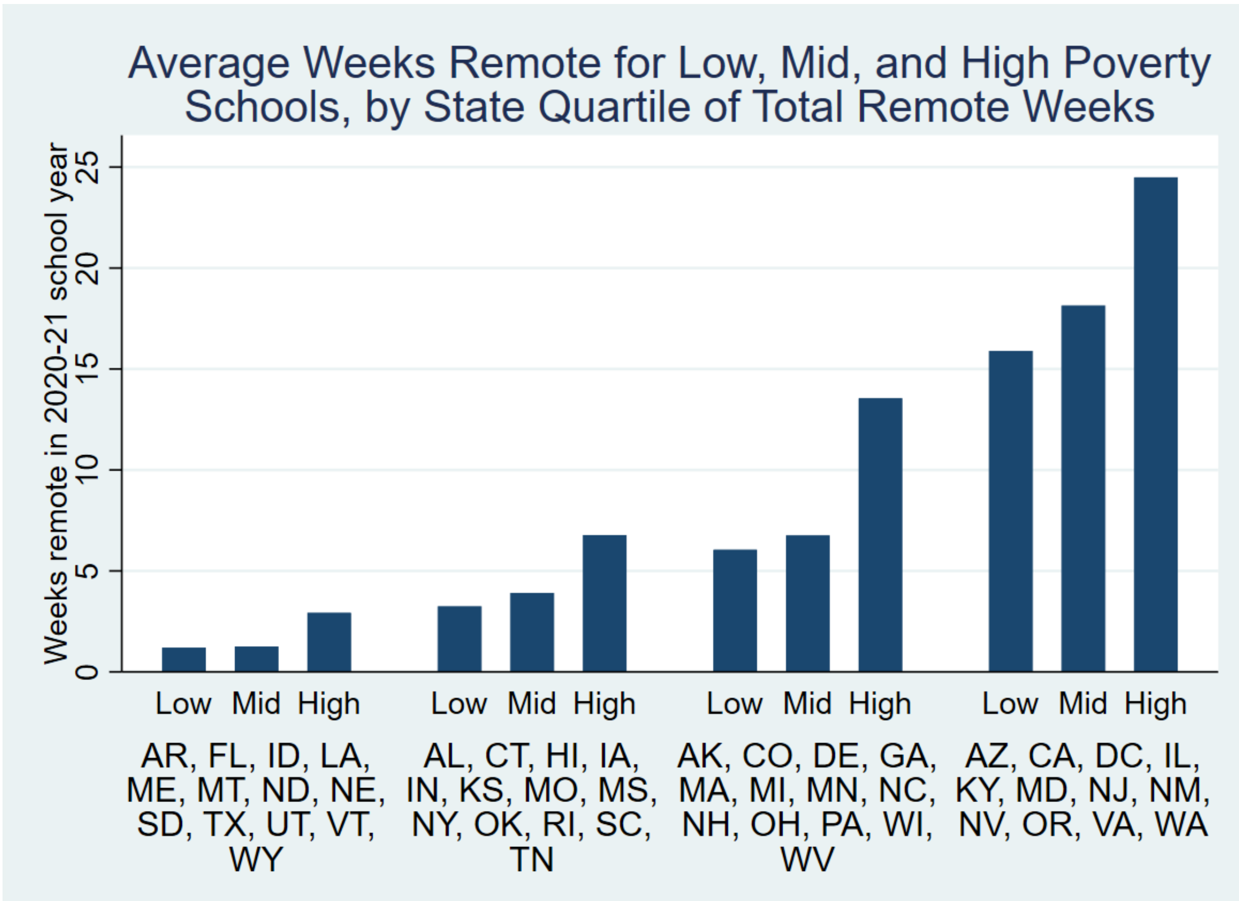
Notes: Sample includes 2,102,909 students in grades 3-8 at the time of their follow-up test. The dependent variable is the difference between a student's standardized 2021 fall NWEA MAP score and their expected score based on baseline characteristics from two years earlier (2019). The parameters for predicting expected scores were drawn from a pre-pandemic regression of fall 2019 scores on baseline characteristics from 2017. Standard errors (clustered at the district level) in parentheses.

Table 2.
Decomposing the Difference in Pandemic Achievement Gains between High and Low Poverty Schools, Math

	Amount	*% of total*
Total Difference Between High and Low Poverty Schools	0.168	100%
Due to Direct Effects of:		
Race	0.014	8%
Baseline Scores	0.016	9%
Conditional Learning Loss in High Poverty Schools That Were Fully in Person	0.002	1%
Due to Differing Incidence of Remote and Hybrid Learning	0.051	30%
Due to Differing Effects of Remote and Hybrid Learning	0.085	51%

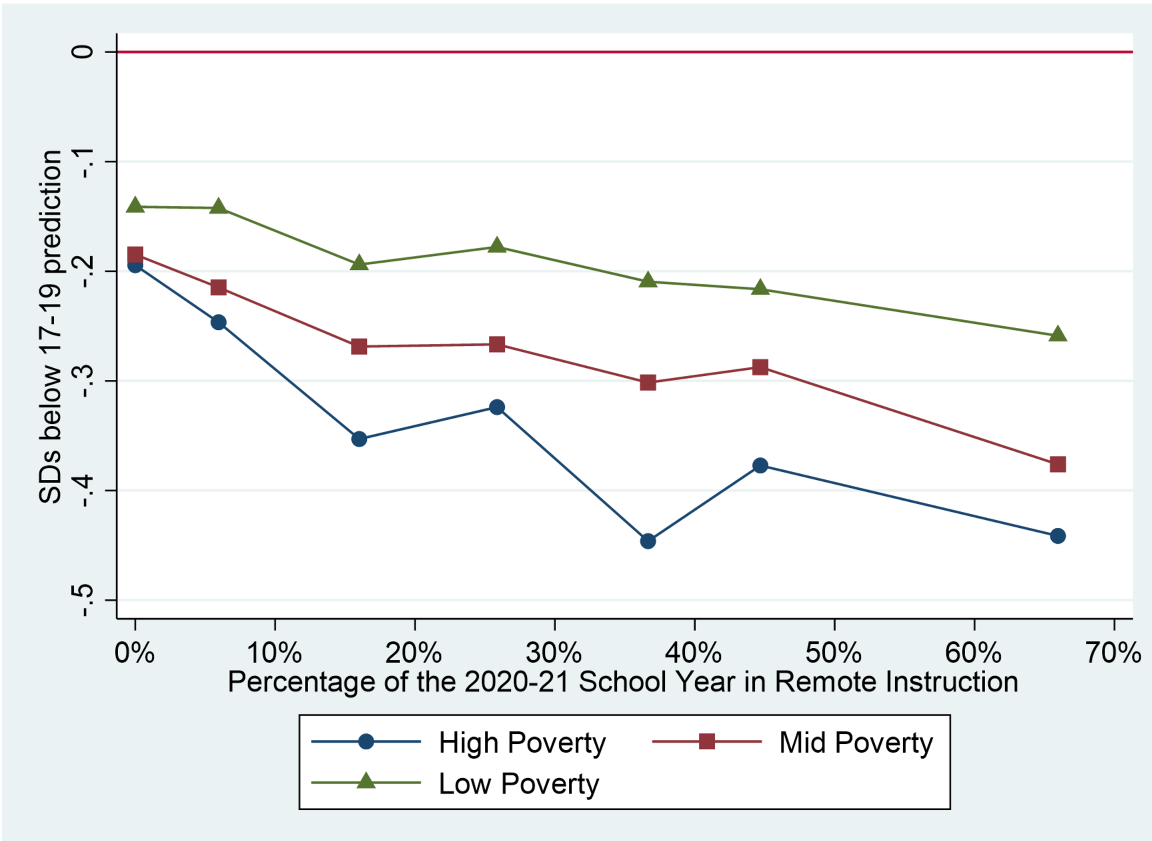
Notes**:** Decomposition based on regression estimates from Table 1, column 5, and based on mean characteristics of high and low poverty schools in the analysis sample used in Table 1. See Appendix B for details on the decomposition and Appendix Table 6 for mean characteristics of high and low poverty schools.

Figure 1. Differences in Remote Instruction by School Poverty Status and State



Note: Weeks of remote instruction are derived from American Enterprise Institute's Return to Learn Tracker. Data on school poverty come from information on the percent of students eligible for free or reduced price lunch (FRPL) in the Common Core Data from 2019-20, or the percentage of students directly certified in the National School Lunch Program if a state did not provide a count of FRPL students. Low poverty schools had fewer than 25 percent of students receiving federal Free or Reduced Price Lunch while high poverty schools had more than 75 percent of students receiving the federal lunch programs.

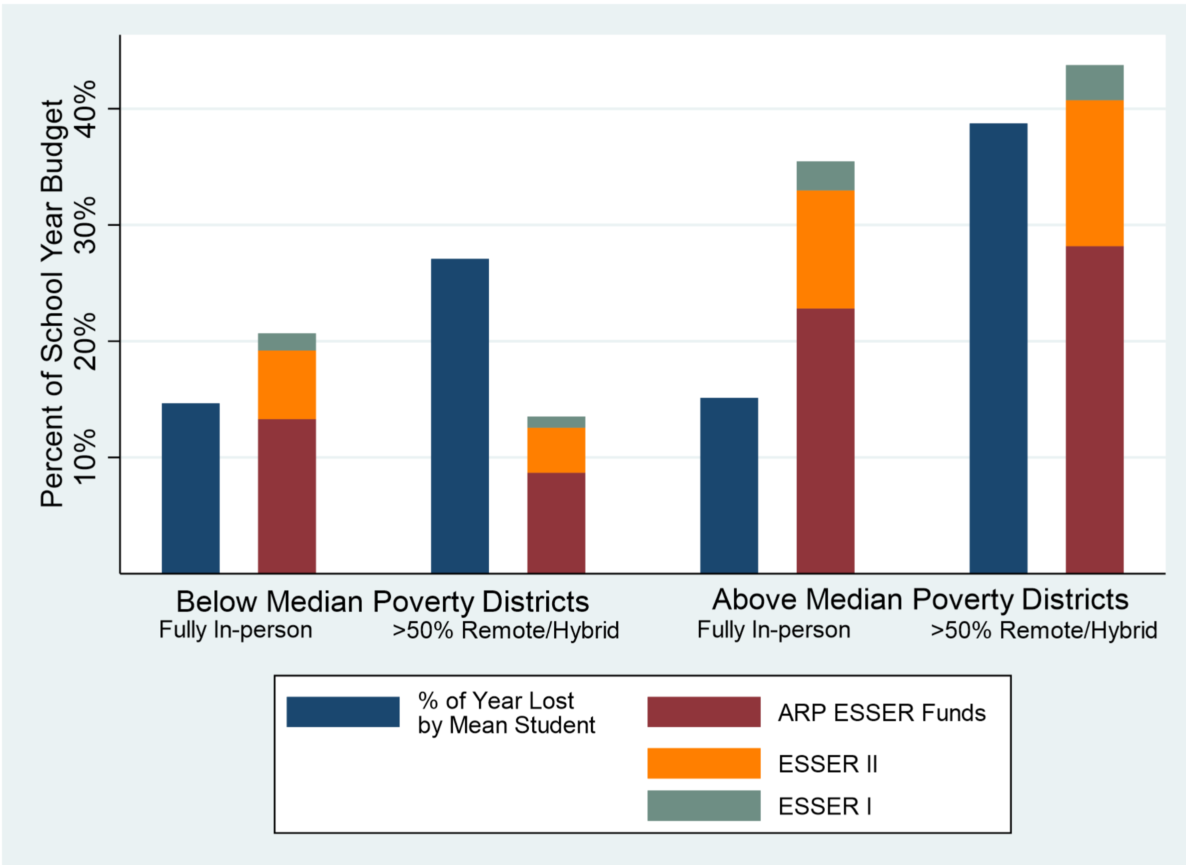
Figure 2. Pandemic Achievement Effects by Remote Schooling and School Poverty, Math



Note: The vertical axis represents the difference between mean fall 2021 achievement and expected achievement based on pre-pandemic growth model estimates. The horizontal axis is the percentage of the 2020-21 school year that a school was in remote instruction. Given the small number of districts that were remote all year, the top category of percent remote combines those who were remote between 50 and 100 percent of the year. Low poverty schools had fewer than 25 percent of students receiving federal Free or Reduced Price Lunch while high poverty schools had more than 75 percent of students receiving the federal lunch programs.

Figure 3.

Pandemic Achievement Losses and Federal Aid as a Share of Annual Spending Math



Note: Achievement effects were converted into weeks of instruction using NWEA growth norms and divided by a 40 week school year (to reflect the fact that salaries and operational expenses are paid by calendar weeks, not the number of instructional weeks in a school year, which is typically 36 weeks). Federal aid is reported relative to the district's annual budget for K-12 schooling, minus capital expenditures. High poverty districts are the half of districts with the highest percent of students receiving Free or Reduced Price Lunch (and low poverty districts are the bottom half). Districts are considered "fully in-person" if the AEI reports no remote or hybrid instruction in the district during the 2020-21 school year.

Appendix A: Explaining the Change in School Effects 2017-19 to 2019-21

To estimate how the variance of school effects changed between the pre-pandemic and pandemic periods, we use a two-step approach. We first estimated the following equation by OLS for 2017-19 and 2019-21:

(1) $\mathrm{S_{ij}} = \beta_0 + X_{ij}\beta_4 + \delta_j + \varepsilon_{ij}$

where X includes all the student-level covariates and δ_j are school fixed effects. We then use the estimated school fixed effects plus the student-level residuals, $\hat{\delta}_j + \hat{\varepsilon}_i = \mathrm{S_{ij}} - \hat{\beta}_0 + X_{ij}\hat{\beta}_4$, as the dependent variable in a simple hierarchical linear model for each year with only an intercept and school random effects, estimated using the xtreg command in Stata. This yields estimates of the variance of the underlying school (σ_μ^2) and student (σ_ε^2) error components in each year. If the pandemic introduced school-level shocks then the variance of school effects will be larger in 2021 than it was in 2019, e.g., $\sigma_{\mu,2021}^2 > \sigma_{\mu,2019}^2$.

We then re-estimated the hierarchical models controlling for three school poverty categories, percent remote and hybrid, and their interactions. If school poverty and remote/hybrid instruction capture the pandemic-related school-level shocks, then the school-level variance estimate from this model should be lower in 2021 compared to a model that does not control for any school characteristics.

As can be seen from the table below, the variance in the school effect rose substantially between 17-19 and 19-21 for both math (.0202, 81% rise) and reading (.0133, 60% rise). Controlling for poverty and hybrid/remote explains little of the school-level variance in 17-19 but explains a much larger proportion of variation in 19-21. Overall, Controlling for poverty and hybrid/remote accounted for 66% of the rise in school-level variance for math, and 57% for reading.

	Math			Reading		
	17-19	19-21	Change	17-19	19-21	Change
Variance of School Effect	0.0248	0.0450	0.0202	0.0220	0.0353	0.0133
Variance of School Effect Controlling for Poverty and Hybrid/Remote	0.0216	0.0283	0.0068	0.0189	0.0247	0.0058
% of Change in School Variance Accounted for by Poverty and Hybrid/Remote:			66%			57%

Appendix B:
Decomposing the Role of Disparate Incidence and Disparate Impacts of Remote/Hybrid instruction on Pandemic Achievement Differences between High and Low Poverty Schools

We use the parameters from Column (5) of Table 1 to identify the share of the widening attributable to multiple factors. Below, the subscript for each coefficient refers to the row number from Table 1.

$$\bar{R}_{Low} - \bar{R}_{Hgh} =$$

$$+\hat{\gamma}_1\left(\overline{Black}_{Low} - \overline{Black}_{Hgh}\right) + \hat{\gamma}_2\left(\overline{Hispanic}_{Low} - \overline{Hispanic}_{Hgh}\right) + \qquad \text{(a)}$$
$$\hat{\gamma}_3\left(\overline{Asian}_{Low} - Asian_{Hgh}\right) + \hat{\gamma}_4\left(\overline{Other}_{Low} - \overline{Other}_{Hgh}\right) +$$
$$\hat{\gamma}_5\left(\overline{MidBase}_{Low} - \overline{MidBase}_{Hgh}\right) + \hat{\gamma}_6\left(\overline{LowBase}_{Low} - \overline{LowBase}_{Hgh}\right)$$

$$-\hat{\gamma}_8 \qquad \text{(b)}$$

$$(\hat{\gamma}_{12} + \hat{\gamma}_{14})\left(\overline{\%Hybrid}_{Low} - \overline{\%Hybrid}_{Hgh}\right) + (\hat{\gamma}_9 + \hat{\gamma}_{11})\left(\overline{\%Remote}_{Low} - \overline{\%Remote}_{Hgh}\right) \qquad \text{(c)}$$

$$-\hat{\gamma}_{14}\left(\overline{\%Hybrid}_{Low}\right) - \hat{\gamma}_{11}\left(\overline{\%Remote}_{Low}\right) \qquad \text{(d)}$$

The first component, (a), captures the differences in student growth due to differences in the race/ethnicity and baseline achievement of students. The second component, (b), reflects the differential losses of high and low-poverty schools that were in person throughout 2020-21. The third component, (c), measures the effect of disparate incidence of remote and hybrid instruction, assessed as the impact of remote and hybrid instruction for high poverty schools. The fourth component, (d), is the largest component. It reflects the differential impact of remote schooling on high poverty schools.

Appendix Table 1:
Pandemic Achievement Gains by Student and School Characteristics, Reading

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Race (Reference: White)							
Black	-0.080		-0.062	-0.023	-0.019	-0.039	-0.018
	(0.010)		(0.008)	(0.004)	(0.007)	(0.005)	(0.007)
Hispanic	-0.066		-0.048	-0.030	-0.007	-0.039	-0.007
	(0.015)		(0.014)	(0.003)	(0.007)	(0.003)	(0.007)
Asian	0.018		0.013	-0.019	0.019	-0.017	0.019
	(0.010)		(0.009)	(0.005)	(0.007)	(0.005)	(0.008)
Other	-0.023		-0.016	-0.011	-0.005	-0.015	-0.005
	(0.008)		(0.008)	(0.003)	(0.007)	(0.004)	(0.007)
Baseline Score (Reference: Top Quartile)							
Middle Quartiles		-0.048	-0.039	-0.013	-0.031	-0.019	-0.031
		(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Bottom Quartile		-0.115	-0.098	-0.043	-0.076	-0.052	-0.076
		(0.010)	(0.008)	(0.005)	(0.006)	(0.006)	(0.006)
School Poverty (Reference: Low <25%)							
Middle (25%-75%)					-0.021	0.019	-0.021
					(0.009)	(0.015)	(0.009)
High (>75%)					-0.038	0.011	-0.037
					(0.016)	(0.019)	(0.016)
Remote Schooling							
% Remote in 2020-21					-0.081	N/A	-0.079
					(0.024)		(0.025)
Interactions:							
• Middle Poverty					-0.034	-0.081	-0.033
					(0.023)	(0.024)	(0.023)
• High Poverty					-0.094	-0.133	-0.096
					(0.046)	(0.033)	(0.046)
Hybrid Schooling							
% Hybrid in 2020-21					0.018	N/A	0.018
					(0.013)		(0.013)
Interactions:							
• Middle Poverty					-0.037	-0.008	-0.036
					(0.014)	(0.021)	(0.015)
• High Poverty					-0.074	-0.047	-0.076
					(0.031)	(0.030)	(0.031)
% Tested in School							0.025
							(0.019)
Constant	-0.093	-0.066	-0.056	N/A	-0.027	N/A	-0.050
	(0.004)	(0.003)	(0.003)		(0.008)		(0.019)
Fixed Effects?	No	No	No	School	No	District	No

Notes: Sample includes 1,666,203 students in grades 3-8 at the time of their follow-up test. Dependent variable is the difference between a student's standardized 2021 fall NWEA MAP score and their expected score based on baseline characteristics from two years earlier (2019). The parameters for predicting expected scores were drawn from a pre-pandemic regression of fall 2019 scores on baseline characteristics from 2017. Standard errors (clustered at the district level) in parentheses.

Appendix Table 2:
Decomposing the Difference in Pandemic Achievement Gains between High and Low Poverty Schools, Reading

	Amount	*% of total*
Total Difference Between High and Low Poverty Schools	0.146	100%
Due to Direct Effects of:		
Race	0.008	5%
Baseline Scores	0.021	14%
Conditional Learning Loss in High Poverty Schools That Were Fully in Person	0.038	26%
Due to Differing Incidence of Remote and Hybrid Learning	0.028	19%
Due to Differing Effects of Remote and Hybrid Learning	0.051	35%

Notes**:** Decomposition based on regression estimates from Appendix Table 1, column 5, and based on mean characteristics of high and low poverty schools in the analysis sample used in Appendix Table 1. See Appendix B for details on the decomposition and Appendix Table 6 for mean characteristics of high and low poverty schools.

Appendix Table 3: Comparing the Analysis Sample to the Universe of K-8 Public Schools

	19-21 Analysis Sample, Math	19-21 Analysis Sample, Reading	CCD Grades 3-8
Race			
White	52%	52%	46%
Black	13%	14%	15%
Hispanic	20%	19%	28%
Asian	4%	4%	5%
Poverty level			
High	22%	22%	27%
Mid	54%	55%	54%
Low	24%	23%	20%
Urbanicity			
City	25%	25%	30%
Rural	19%	20%	20%
Suburb	44%	43%	39%
Town	12%	12%	11%
Learning Mode			
Mean % of Year Remote	21%	20%	24%
Mean % of Year Hybrid	47%	47%	46%
Mean NWEA Fall 2021 Normalized RIT Score	-0.11	-0.08	N/A
Number of Schools in Sample	9,692	9,490	74,189
Number of Students in Sample	2,102,909	1,666,203	22,835,038

Notes: Analysis samples include students in NWEA test score data that (1) attend schools that test at least 10 students in fall 2017, fall 2019, and fall 2021; (2) attend schools that test at least 60% of their school-grade-level enrollment as reported in the Common Core of Data; and (3) have available data on the student's race, gender, school poverty level, and learning modality.

Appendix Table 4: 2017-19 Growth Model Parameters

	Math	Reading
Race (Reference: White)		
Black	-0.116	-0.112
	(0.006)	(0.006)
Hispanic	-0.024	-0.028
	(0.005)	(0.005)
Asian	0.195	0.136
	(0.007)	(0.006)
Other	-0.028	-0.033
	(0.005)	(0.006)
School Poverty (Reference: Low <25%)		
Middle (25%-75%)	-0.082	-0.077
	(0.010)	(0.011)
High (>75%)	-0.175	-0.142
	(0.016)	(0.015)
Linear Term of Baseline Score	0.757	0.729
	(0.004)	(0.005)
Remote Schooling		
% Remote in 2020-21	0.044	0.035
	(0.035)	(0.024)
Interactions:		
• Middle Poverty	-0.038	-0.015
	(0.028)	(0.022)
• High Poverty	-0.049	-0.075
	(0.031)	(0.025)
Hybrid Schooling		
% Hybrid in 2020-21	-0.007	-0.011
	(0.013)	(0.013)
Interactions:		
• Middle Poverty	-0.006	0.002
	(0.014)	(0.014)
• High Poverty	0.054	0.028
	(0.028)	(0.027)
All X's	Yes	Yes
School FE	No	No
District FE	No	No

Notes: Sample includes 2,313,927 students in math and 1,822,756 students in reading in grades 3-8. Dependent variable is the student's fall 2019 test score. The parameters for predicting expected scores in Table 1 and Appendix Table 4 are drawn from these regressions. Standard errors (clustered at the district level) in parentheses.

Appendix Table 5: Predictors of Having a Follow-up Score

	2017-19	2019-21

	Math	Reading	Math	Reading
Race (Reference: White)				
Black	-0.080	-0.055	-0.075	-0.049
	(0.029)	(0.027)	(0.015)	(0.016)
Hispanic	-0.016	-0.016	-0.001	0.010
	(0.013)	(0.014)	(0.014)	(0.019)
Asian	-0.061	-0.049	-0.045	-0.010
	(0.014)	(0.013)	(0.010)	(0.016)
Other	-0.039	-0.044	-0.046	-0.045
	(0.011)	(0.009)	(0.015)	(0.013)
School Poverty (Reference: Low <25%)				
Middle (25%-75%)	-0.054	-0.036	-0.060	-0.073
	(0.024)	(0.026)	(0.025)	(0.025)
High (>75%)	-0.073	-0.044	-0.024	-0.030
	(0.030)	(0.032)	(0.028)	(0.029)
Linear Term of Baseline Score	0.014	-0.010	0.006	-0.039
	(0.004)	(0.007)	(0.004)	(0.011)
Remote Schooling				
% Remote in 2020-21	-0.069	-0.106	-0.235	-0.304
	(0.056)	(0.062)	(0.060)	(0.084)
Interactions:				
• Middle Poverty	0.012	0.002	0.212	0.118
	(0.070)	(0.067)	(0.047)	(0.094)
• High Poverty	-0.002	0.017	0.087	-0.022
	(0.077)	(0.074)	(0.092)	(0.098)
Hybrid Schooling				
% Hybrid in 2020-21	-0.016	-0.015	0.020	-0.027
	(0.042)	(0.043)	(0.025)	(0.028)
Interactions:				
• Middle Poverty	0.088	0.063	0.018	0.061
	(0.045)	(0.046)	(0.034)	(0.035)
• High Poverty	0.109	0.087	-0.063	0.000
	(0.058)	(0.057)	(0.047)	(0.052)
All X's	Yes	Yes	Yes	Yes
School FE	No	No	No	No
District FE	No	No	No	No

Notes: Sample includes all students in grades 1-6 with a baseline score and non-missing independent variables. Dependent variable is whether the student had a follow-up score in either fall 2019 (in the 2017-19 regressions) or fall 2021 (in the 2019-21 regressions). Standard errors (clustered at the district level) in parentheses.

Appendix Table 6: Mean Student Characteristics by School Poverty

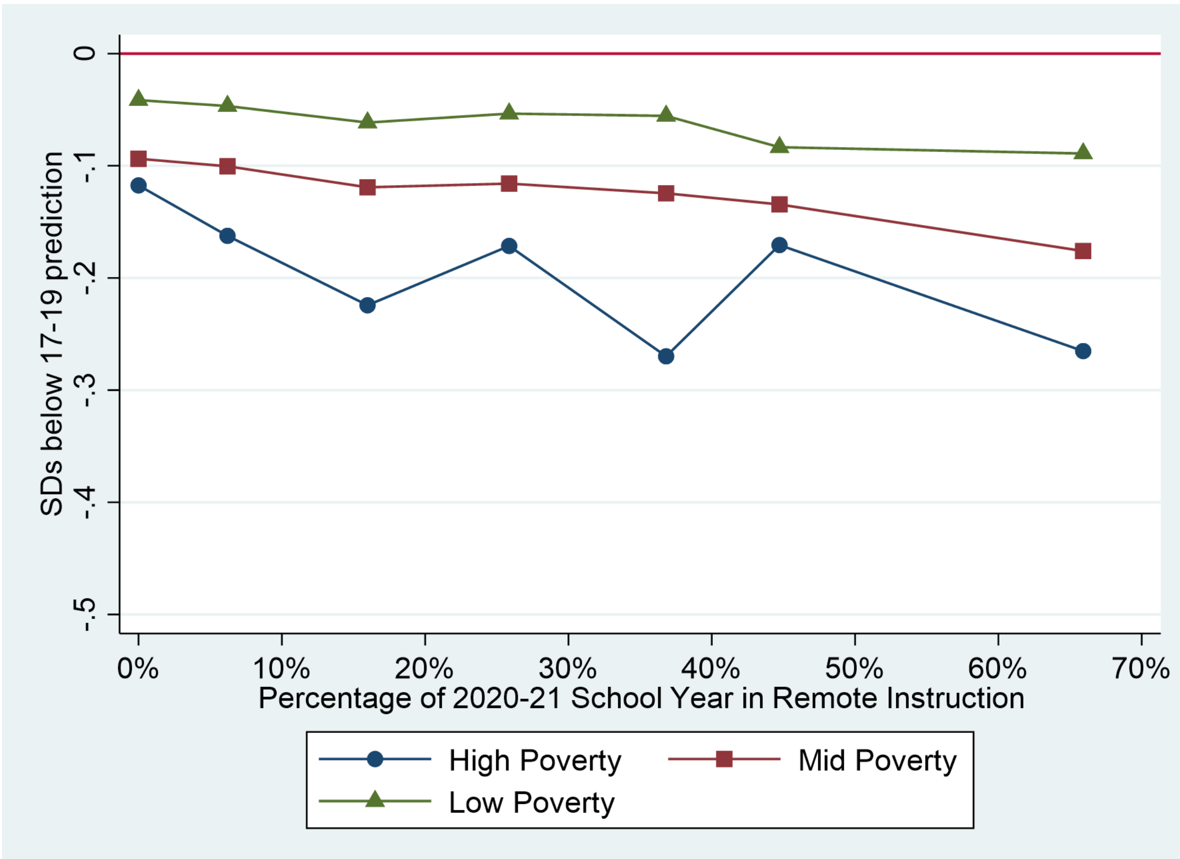
	Math	Reading

	Low Poverty	High Poverty	Low Poverty	High Poverty
Race				
White	68.7%	22.0%	70.0%	23.2%
Black	4.2%	27.0%	4.4%	29.0%
Hispanic	7.4%	40.1%	7.4%	36.8%
Asian	8.0%	2.3%	7.6%	2.2%
Other	11.7%	8.6%	10.6%	8.8%
Baseline score				
High	41.5%	11.4%	40.1%	11.8%
Mid	46.8%	47.8%	47.2%	48.2%
Low	11.7%	40.8%	12.7%	40.0%
% of 2020-21 Remote	14.7%	33.5%	13.4%	32.1%
% of 2020-21 Hybrid	53.0%	42.0%	52.4%	43.3%

Note: These means are used for the decomposition calculation presented in Table 2 and Appendix B.

Appendix Figure 1.

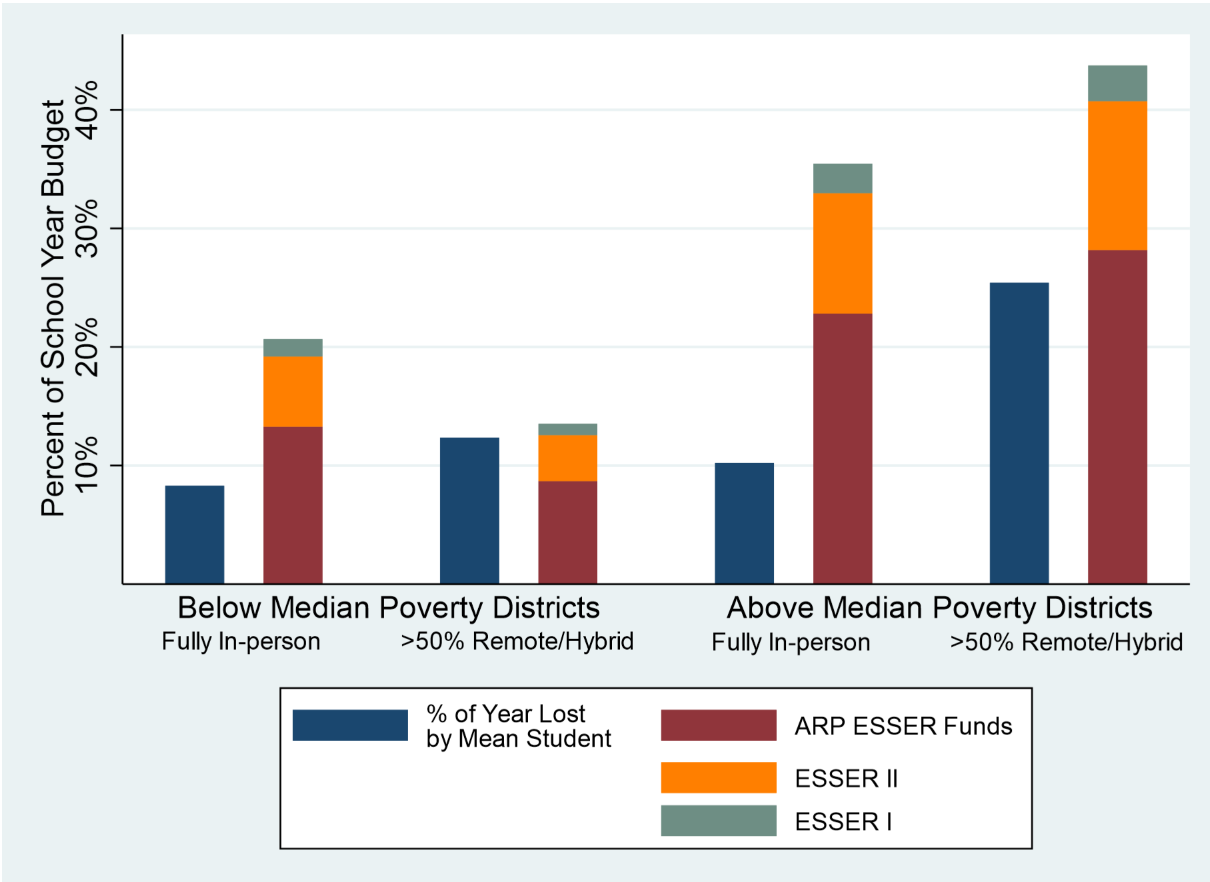
Pandemic Achievement Effects by Remote Schooling and School Poverty, Reading



Note: The vertical axis represents the difference between mean fall 2021 achievement and expected achievement based on pre-pandemic growth model estimates. The horizontal axis is the percentage of the 2020-21 school year that a school was in remote instruction. Given the small number of districts that were remote all year, the top category of percent remote combines those who were remote between 50 and 100 percent of the year. Low poverty schools had fewer than 25 percent of students receiving federal Free or Reduced Price Lunch while high poverty schools had more than 75 percent of students receiving the federal lunch programs.

Appendix Figure 2.

Pandemic Achievement Losses and Federal Aid as a Share of Annual Spending, Reading



Note: Achievement effects were converted into weeks of instruction using NWEA growth norms and divided by a 40 week school year (to reflect the fact that salaries and operational expenses are paid by calendar weeks, not the number of instructional weeks in a school year, which is typically 36 weeks). Federal aid is reported relative to the district's annual budget for K-12 schooling, minus capital expenditures. High poverty districts are the half of districts with the highest percent of students receiving Free or Reduced Price Lunch (and low poverty districts are the bottom half). Districts are considered "fully in-person" if the AEI reports no remote or hybrid instruction in the district during the 2020-21 school year.