Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1571

Post-pandemic achievement recovery based on student characteristics and district-level policies

Jessica Sperber Knott
Max Planck Institute for Human Development

Yu Bai
Duke University

Jade M. Jenkins
University of California, Irvine

Kenneth A. Dodge
Duke University

Tyler W. Watts
Teachers College, Columbia University

It is well-established that student achievement has declined dramatically since the onset of the COVID-19 pandemic. However, less is known how these patterns may vary by developmental timing of exposure to the pandemic and how district-level policies may have influenced student recovery. This descriptive study examined patterns of achievement across middle childhood during and after the COVID-19 pandemic for the population of Grade 3 through 8 students attending North Carolina public schools, and the role of pandemic relief policy in mitigating learning loss (2017-2023; N=1,463,253, 51% Male, 25% Black, 20% Hispanic, 45% White). Students showed evidence of recovery in academic achievement, though scores still lag behind pre-pandemic benchmarks. Low-income students demonstrated strong evidence of recovery in reading. Younger students (Grade 3 in 2021) demonstrated the strongest rates of recovery across subjects, suggesting important variation based on the developmental timing of exposure to the pandemic. Finally, district-level policies of in-person instruction and pandemic relief funding explained only modest variation in achievement scores above and beyond pre-pandemic levels of achievement. These findings contribute to a growing body of pandemic-era literature by examining heterogeneity in achievement recovery by developmental timing of exposure and district-level policies.

VERSION: September 2026

Suggested citation: Sperber Knott, Jessica, Yu Bai, Jade M. Jenkins, Kenneth A. Dodge, and Tyler W. Watts. (2026). Post-pandemic achievement recovery based on student characteristics and district-level policies. (EdWorkingPaper: 26-1571). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/em8a-nw26

Post-pandemic achievement recovery based on student characteristics and district-level policies

Jessica Sperber Knott[1], Yu Bai[2], Jade V. M. Jenkins[3], Kenneth Dodge[2], Tyler W. Watts[4]

[1] Max Planck Institute for Human Development
[2] Duke University
[3] University of California, Irvine
[4] Teachers College, Columbia University

Acknowledgements: This publication was supported by grant funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (5R01HD095930-07 to JVMJ). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The materials, data, and analytic code necessary to reproduce the analyses presented here are not publicly accessible. The analyses presented here were not pre-registered. The authors report no conflicts of interest.

Corresponding Author:
Tyler Watts
525 W 120th St.
New York, NY 10027
tww2108@tc.columbia.edu

ABSTRACT: It is well-established that student achievement has declined dramatically since the onset of the COVID-19 pandemic. However, less is known how these patterns may vary by developmental timing of exposure to the pandemic and how district-level policies may have influenced student recovery. This descriptive study examined patterns of achievement across middle childhood during and after the COVID-19 pandemic for the population of Grade 3 through 8 students attending North Carolina public schools, and the role of pandemic relief policy in mitigating learning loss (2017-2023; N=1,463,253, 51% Male, 25% Black, 20% Hispanic, 45% White). Students showed evidence of recovery in academic achievement, though scores still lag behind pre-pandemic benchmarks. Low-income students demonstrated strong evidence of recovery in reading. Younger students (Grade 3 in 2021) demonstrated the strongest rates of recovery across subjects, suggesting important variation based on the developmental timing of exposure to the pandemic. Finally, district-level policies of in-person instruction and pandemic relief funding explained only modest variation in achievement scores above and beyond pre-pandemic levels of achievement. These findings contribute to a growing body of pandemic-era literature by examining heterogeneity in achievement recovery by developmental timing of exposure and district-level policies.

PUBLIC SIGNIFICANCE STATEMENT: Although students still lag behind their pre-pandemic peers in academic achievement by the end of 2023, they have also demonstrated strong evidence of recovery. However, these patterns were not equal across students: younger students tended to ‘catch-up’ more quickly than older students, and students engaged in more in-person learning during lockdown tended to perform better on standardized tests at the end of 2021.

Keywords: COVID-19, longitudinal, middle childhood, academic trajectories, developmental processes

The immediate negative impact of the COVID-19 pandemic on children's academic achievement has been well-documented over recent years. A recent meta-analysis concluded that, on average, students lost about 0.19 SDs in achievement from pre-pandemic levels, with the greatest impacts detected for math and science (Di Pietro, 2023). Several reports have documented disparate impacts for economically disadvantaged students, exacerbating previously established achievement gaps between higher- and lower-income students (Dewey et al., 2025; Kuhfeld, Soland, & Lewis, 2022; Kuhfeld, Soland, Lewis, et al., 2022; Sass & Goldring, 2021).

In the years since the pandemic, the recovery literature has generally concluded that while students demonstrated initial learning boosts after school reopening, test scores continue to lag behind pre-pandemic benchmarks. Annual reports by the Northwest Evaluation Association (NWEA) – a large test assessment provider in the United States – reported accelerated rates of learning in school year (SY) 2021-2022 (Kuhfeld & Lewis, 2022) but stalled progress toward recovery in SY 2022-2023 (Lewis & Kuhfeld, 2023). An extension of this analysis using multiple types of assessments through SY 2024 indicated that these lags are particularly pronounced for English language arts (Kuhfeld & Lewis, 2025). The administrators of the *i-Ready* math and reading assessment reported similarly slow rates of rebounding and stalled progress toward recovery (Curriculum Associates, 2023; 2025). Additionally, a nationwide analysis of state summative assessments found continued accelerated learning across the first few years of school reopening, but growing gaps in reading achievement in 2024 (Fahle et al., 2024).

While numerous studies have examined differences in these patterns by economic disadvantage, much of this work has relied on aggregated school-level or district-level data (e.g,. Kuhfeld et al., 2023; Dewey et al., 2024). For the few datasets by which student-level data is available (e.g., NAEP), there is a lack of longitudinal data to assess change in these patterns over

time. In addition, there remains conflicting accounts among the few studies that have explicitly tested variation in these patterns by developmental timing of exposure (Curriculum Associates, 2025; Lewis & Kuhfeld, 2024). This is an important area of interest for developmental scientists, as decades of research have documented non-linear patterns of growth in academic achievement across childhood (e.g., Cameron et al., 2015; Mok et al., 2015), and developmental theory suggests that younger children may be more sensitive to environmental inputs than older children. Thus, we might expect to observe important variation in pandemic recovery by child age. Finally, there is great interest in understanding what types of COVID-era policies may have boosted or hindered achievement recovery. Several studies have attempted to generate causal estimates of the effect of pandemic relief funding and in-person instruction on student achievement outcomes (e.g., Goldhaber & Falken, 2025; Goldhaber et al., 2023). Obtaining causal estimates of such policies is difficult due to the non-random nature of their implementation. However, the use of longitudinal achievement data across the pre-pandemic and post-pandemic periods is useful for improving estimates, as it can remove substantial confounding variation that might bias analyses relying solely on cross-sectional comparisons.

Our study provides descriptive evidence regarding COVID-era achievement in a large and diverse state: North Carolina. Our data contains information on the entire population of public-school students in the state, and we used this administrative database to examine basic questions about achievement loss and recovery, moderating factors, and how COVID-era policies may have shaped these patterns. First, we examined state-wide trends of post-pandemic achievement for all children in Grades 3 through 8, and how these trends differed by socioeconomic status. We also examined how the developmental timing of exposure to the pandemic may have shaped initial achievement losses and recovery rates.

After examining these patterns using child-level characteristics, we then examined recovery patterns by district-level policies. We leveraged open-source data on the amount of in-person instruction and pandemic relief funding allocated to each school district to examine the association between each policy and COVID recovery. This analysis is restricted to a longitudinal sub-sample of students with available data from both the pre- and post-pandemic periods. This allows for the estimation of these associations while adjusting for each child's initial levels of achievement prior to the pandemic. Though not a causal estimate, this approach allowed us to compare the effects of a policy while holding constant prior levels of achievement, improving causal inference.

As such, our paper extends the COVID literature in several ways. We analyze population-level data from a diverse state using a novel assessment of achievement that has not yet been examined within the pandemic-era literature. Our approach also applies a developmental lens to these questions by considering how child age may explain important variation in achievement outcomes, deepening the field's understanding of how environmental shocks shape children's initial response and subsequent learning in the years after schools reopened. Finally, we provide new evidence on the association between COVID-era policies and achievement using a longitudinal sub-sample assessed across the pre-pandemic and post-pandemic period.

Developmental Timing

When considering the factors that could influence student learning during the pandemic, the developmental timing of exposure to the COVID-19 lockdowns could constitute a key source of variation. Younger students may have had more access to instructional content that could substitute for the lost time in school because early literacy and numeracy activities can be more easily supported at home. When students returned to school, they may have had more chances to

"catch-up" because early-grade instruction includes substantial redundancy (Engel et al., 2013). On the other hand, one might expect older students to fare better during the pandemic because older children are more likely to have the self-regulatory and metacognitive skills necessary to engage in virtual learning, whereas a young child may require additional support to stay on task and engaged.

In an examination of 5 million students across Grades 3 – 8, Kuhfeld and colleagues (2023) found that students in Grades 3 through 5 demonstrated the largest declines in reading skills immediately after school reopening in 2021, whereas declines in math performance were more uniform across the age distribution. This pattern suggests that younger students' reading abilities may be particularly sensitive to in-school instruction. In a separate examination of the same sample, Peters and colleagues (2025) found low-achieving students demonstrated worse recovery in the post-pandemic period than higher-achieving peers, with little variation by student grade. In contrast, Sass and Ali (2024) found that students in Grades 1 – 3 during the pandemic fared worse in math performance upon the return to school compared to students who were in Grades 5 – 7 at the pandemic's onset. Greater longitudinal research is necessary to clarify patterns by student age for the average student, in addition to those in the tails of the distribution.

Academic Achievement and District-Level Policies

Like all school systems, North Carolina's education system was severely disrupted by the COVID-19 pandemic. Given distributed decision-making power across 115 public school districts, the state witnessed a wide range of district-level responses to school closures and reopening decisions. Indeed, districts made policy decisions for many schools throughout the pandemic response, including in-person instruction procedures and spending pandemic relief funding (Christian et al., 2022; Troppe et al., 2024). For example, after the closure of schools in

April 2020, only 35% of districts required synchronous (real-time) teaching for all grade levels (Osborne, Rash, et al., 2020). During the fall of 2020, about 45% of districts adopted hybrid learning policies for all or some student groups, while 55% continued with fully remote instruction (Osborne, Clifford, et al., 2020). Notably, districts with a large population of students of color were more likely to opt for fully remote learning (76%) over hybrid procedures (24%; Osborne, Clifford, et al., 2020). In contrast, this propensity was reversed for districts serving predominantly White students, with three-quarters (76%) opting for hybrid learning over remote procedures (24%; Osborne, Clifford, et al., 2020).

Understanding the role of such policies in mitigating or exacerbating the adverse effects of the COVID-19 pandemic on children's achievement remains important for the field. Indeed, several correlational studies from across the U.S. have reported associations between greater remote instruction and greater learning losses upon the return to school. However, it is challenging to generate causal estimates of these policies due to the non-random nature of their implementation. For example, the enactment of hybrid learning procedures in schools was deeply confounded by political, regional, and demographic factors. One natural experiment design from The Netherlands identified a 0.08 SDs loss in achievement as a result of an 8-week, nation-wide period of remote instruction (Engzell et al., 2021).

Setting aside endogeneity concerns, capturing the association between district policies and district-level achievement variation may reveal valuable insights about local stakeholders' ability to influence student outcomes during a disruptive period. For example, in a case study of three metro-Atlanta school districts, Sass and Goldring (2021) reported large variation across districts in achievement declines and growth rates after the pandemic. While district-level factors can be biased due to self-selection into districts, examining district-level variation can reveal an

"upper bound" on the extent to which district-level differences might have affected both initial drops in achievement and post-COVID recovery rates. This is particularly interesting to study in North Carolina, a large yet demographically and politically diverse state within the United States.

The Present Study

The present study leveraged administrative data to describe patterns of achievement losses and recovery on a summative assessment of reading and math skills since the onset of the COVID-19 pandemic in North Carolina. First, we compared post-pandemic achievement scores (2021-2023) to pre-pandemic benchmarks (2017-2019) in various cohorts of students. We also explored heterogeneity in these patterns by child age and socioeconomic disadvantage. Next, we examined a sub-sample of students with data available longitudinally across both the pre-pandemic and post-pandemic periods to investigate if district-level policies were associated with post-pandemic achievement and if they explained variation in achievement across districts.

Method

Data

We leveraged population-level administrative data managed by the North Carolina Educational Research Center, which holds the administrative records for all children in Grades 3 – 8 enrolled in the North Carolina Public School System. The term School Year (SY) refers to the period from July 1 of the preceding year to June 30 of the current year. The dataset contains the annual End-of-Grade test scores of all children enrolled in North Carolina's public schools between SY 2017 – 2019 and SY 2021 – 2023; there were no summative achievement tests administered in SY 2020. Analyses were restricted to students with a non-missing reading score in Grades 3 through 8 enrolled in one of North Carolina's 115 public school districts. Due to a change in math test administration procedures for students in Grade 8, analyses using math

scores were restricted to Grades 3 through 7. The full dataset consisted of 1,463,253 unique students, with 3,729,855 individual records.

Sample

Due to our interest in examining both post-pandemic recovery at the child level and COVID-era policies at the district level, we analyzed two subsets of this larger dataset. First, we restricted the sample to students with at least three successive reading or math observations in the pre-pandemic or post-pandemic period observed within the same district. This allowed us to compare pre-pandemic versus post-pandemic achievement scores within the same cohort of children over time. The resulting sample consisted of 363,580 individual records from the pre-pandemic period (2017-2019), and 299,693 individual records from the post-pandemic period (2021-2023). This cross-sectional subsample of students is demographically similar to the full sample of students enrolled North Carolina schools during the same period (Supplemental Table 1). Table 1 compares the demographic characteristics of these longitudinal student cohorts over time (described in greater detail below).

To examine district-level variation in achievement based on COVID-era policies, we further restricted the sample to a longitudinal cohort of children who were observed both before and after the onset of the pandemic. This subsample consisted of 73,289 individual children with valid math and reading scores observed in Grade 3 (2019), Grade 5 (2021), Grade 6 (2022), and Grade 7 (2023). This sample represented 113 out of 115 North Carolina school districts. We chose to examine this longitudinal cohort as it allowed us to adjust for pre-pandemic measures of achievement, further removing variation that may confound the associations between pandemic-era policies and student achievement. This longitudinal sub-sample was not substantively

different from the full sample of children enrolled in North Carolina schools, or from those included in cross-sectional analyses (Supplemental Table 1).

It is possible that the pandemic changed the landscape of students who enrolled in North Carolina public schools. Therefore, we examined demographic differences in each cohort across the pre-pandemic and post-pandemic periods using t-tests and included these characteristics as covariates in all of our primary analyses (Table 1). We also conducted logistic regressions to examine which student characteristics were predictive of children re-enrolling in North Carolina public schools in SY 2021 (among students who were enrolled prior to the pandemic). Results are presented as odds ratios in Supplemental Table 2. Students who had ever repeated a grade before the pandemic were most likely to attrit from the sample (OR = 1.20, 95% CI [1.09, 1.26], $p < 0.001$). Importantly, we did not observe substantial differences in post-COVID attrition based on reading (OR = 1.00, 95% C.I. [1.00, 1.00], $p < 0.001$) or math scores in SY 2019 (OR = 0.99, 95% C.I. [0.99, 0.99], $p < 0.001$).

Student Cohorts

To compare learning growth across middle childhood periods, students were assigned to longitudinal cohorts which reflect the expected grade progression over the course of three years (i.e., Cohort 1 = Grades 3, 4, and 5; Cohort 2 = Grades 4, 5, and 6; Cohort 3 = Grades 5, 6, and 7; Cohort 4 = Grades 6, 7, and 8). These cohorts were generated for both pre-pandemic (2017-2019) and post-pandemic observations (2021-2023) and allowed us to generate simple comparisons in achievement across these two eras. For example, in the youngest cohort of students, the achievement of students enrolled in Grade 3 in 2017 was directly compared to the achievement of students enrolled in Grade 3 in 2021. These same cohorts of students were then compared again as they progressed to Grade 4 (2018 vs. 2022) and Grade 5 (2019 vs. 2023).

Table 1 presents each cohort's demographic characteristics by whether they were observed before or after the pandemic.

Measures

Academic Achievement

Students in North Carolina complete a summative End-of-Grade assessment of reading and math skills at the end of the Spring semester administered by the North Carolina Department of Instruction (NCDPI) across all grades beginning in Grade 3. These assessment scores were transformed into Lexile (reading) and Quantile (math) scores, a framework developed by MetaMetrics that places student ability and content complexity on a common scale spanning grades and years (Sanford-Moore, 2022). These scores allow for within and between-test comparisons to provide a unified framework for describing children's abilities across different types of assessments and variable test content, and scores are linearly scaled to reflect growth over time (Sanford-Moore, 2022). As demonstrated by Figure 1, students' average Lexile and Quantile scores predictably increased with grade and content complexity.

Importantly, NCDPI implemented an updated edition of the End-of-Grade reading and math tests in SY 2019, which continued to be administered through SY 2023 (except in SY 2020, when no test was administered). While the Lexile/Quantile scales are designed to be adaptive to test content changes, we nonetheless observed notable differences in children's Quantile scores in 2019 (Lexile scores did not appear to be affected by this test change; Figure 1). This change in average Quantile scores affects our ability to confidently examine change in math achievement over time. Consequently, we compare pre-pandemic and post-pandemic reading skills in the main text, and reserve math skills for the Supplement.

Post-COVID Indicator

A dichotomous indicator was generated to identify if a test score was observed before (0 = SY 2017 – 2019) or after COVID-related shutdowns (1 = SY 2021 – 2023). This indicator was used to compare achievement scores across these two periods within each cohort of students.

In-person Instruction Index

NCDPI provided monthly updates on the amount of time schools spent in in-person, hybrid, or remote instruction during the 2020-2021 school year. This learning model data was aggregated and coded by the COVID-19 School Data Hub such that in-person instruction was defined as all or most students receiving in-person instruction 5 days a week; hybrid instruction as a blend of in-person and online activities (which may include synchronous or asynchronous instruction); and fully remote learning (*COVID-19 School Data Hub*, n.d.). A district's predominant learning model for each month was defined as the instructional model most often provided across all schools during that month (COVID-19 School Data Hub Team, 2021). Using this information, we generated a weighted index to reflect the amount of time each district spent in in-person instruction during the 2020 – 2021 school year. We coded the variable by multiplying the number of in-person months by "1", the number of hybrid months by "0.5," and the number of fully remote months by "0." The index has a theoretical range of 0 – 10.

Pandemic Relief Funding

Data on the amount of funding allocated per school district for pandemic relief were downloaded from publicly available websites via NCDPI (*COVID Funds | NCDPI*, n.d.). This data includes all dollars allocated across several federal and state-level initiatives, including the Elementary and Secondary School Emergency Relief (ESSER) allocated via the CARES Act and American Rescue Plan (ARP). A per pupil expenditure for each year (i.e., 2021, 2022, and 2023,

respectively) was generated by dividing the allocated dollar amount by each district's total enrollment. The variable was scaled in hundreds of dollars.

Covariates

Student-level demographic covariates included the following: child sex, race (dummy coded as non-Hispanic Black, Hispanic, and non-Hispanic Other; non-Hispanic White is the reference group), whether a student has ever repeated a grade, whether a student receives special education services, whether the child was ever classified as an English language learner, and an indicator of whether the student is classified as economically disadvantaged (eligibility for free or reduced-price lunch).

District-level covariates were obtained using publicly available data distributed for each fiscal year by the National Center for Education Statistics (NCES; Common Core of Data, n.d.). These included an indicator for whether a school district is predominantly rural; student membership (number of students enrolled in the district each school year); and student-teacher ratio (number of teachers employed by the district divided by student membership). Finally, we adjusted for the reported number of per-pupil expenditures spent by each district in each school year. This variable was generated by summing across all local, state, and federal dollars allocated for each school district and dividing by district student membership.

Analytic Plan

Cross-sectional Analyses: Recovery, Moderation by SES, and Developmental Timing of Exposure

We leveraged a mix of cross-sectional and longitudinal modeling to answer our research questions. Our first aim was to describe student achievement loss and recovery. To achieve this, we implemented a cohort comparison approach, in which we compared pre- and post-pandemic cohorts of students across 3 years of learning. For this analysis, achievement scores were

regressed onto a Post-COVID indicator for each grade within each cohort. For example, "Cohort 1" tracks students as they progress across Grades 3, 4, and 5. When examining pandemic differences for Cohort 1, we compared G3 scores in 2017 to G3 scores in 2021; G4 scores in 2018 to G4 scores in 2022; G5 scores in 2019 to G5 scores in 2023. Next, we included covariates for student-level characteristics as we sought to observe the change in average test scores for economically disadvantaged students. Finally, we introduced covariates for district-level characteristics and a Post-COVID X Low SES interaction term to test whether low-income students were performing better or worse after the pandemic. To understand the influence of developmental timing, we observed patterns across cohorts. In all of these cohort comparison analyses, a random intercept for district was included to account for the nested structure of the data.

Longitudinal Analyses: COVID-era Policies and District-level Achievement Variation

For our second aim, we sought to examine the association between pandemic-era policies and student achievement. We recognized that the raw association among these variables was unlikely to capture exogenous variation and likely reflected factors that are endogenous to children and families within school districts (i.e., selection into districts). Therefore, we restricted these analyses to a longitudinal cohort with data across both the pre-pandemic and post-pandemic period. This restriction resulted in one cohort of students who were enrolled in G3 in 2019, had no G4 scores in 2020 due to COVID, then had in-school assessments in G5 in 2021, G6 in 2022, and G7 in 2023. This design allowed us to control for pre-pandemic achievement levels, removing a substantial amount of variance that is likely related to selection into a given school district. Although still a correlational analysis subject to typical concerns of omitted

variables bias, we consider this an improved estimate over models that rely on comparisons with no pre-test data.

For these analyses, we began with a model that regressed a given post-COVID test score on G3 test scores (i.e., pre-COVID) and student demographics. We again included a random intercept for district. In addition to accounting for the nested structure of the data, this random intercept also allowed us to examine the standard deviation of post-pandemic achievement scores across districts. Next, the in-person index, pandemic relief dollars, and student-level covariates were entered into the model. This model provided both the direct association between these policies and achievement, but also allowed us to observe how much district-level variance in achievement was accounted for by controlling for these policies. Finally, we added district-level covariates to examine if any additional variance in achievement could be accounted by other district-level factors.

Transparency and Openness

Achievement scores were rescaled to improve model convergence in line with previous research by dividing by 100 (see Francis et al., 2018), and the resulting coefficients were ultimately converted back to raw units for ease of interpretation. Analyses were conducted in Stata MP v.19.5. Due to confidentiality reasons and data access agreements, the data and code necessary for reproduction of these findings are not publicly available. This study's design and analysis were not preregistered.

Results

Sample Descriptives and Correlations

As expected, we observed differences in student demographic characteristics between the pre-pandemic and post-pandemic cohorts. Most notably, we observed a ~5% reduction in the

low-income student population after the pandemic compared to pre-pandemic cohorts (*p's* < 0.001; Table 1). However, we did not observe large differences in achievement scores between low-income students who attritted from the sample and those who did not attrit for either reading (OR = 1.00, 95% C.I. [1.00, 1.00], $p = 0.009$) or math (OR = 1.00, 95% C.I. [0.99, 1.00], $p = 0.381$). On average, students spent 3.66 months (SD = 1.38; range 1 – 9) of SY 2021 in remote instruction and were allocated $533 of pandemic relief funding in 2021 (SD = $213), $1,464 in 2022 (SD = $382), and $1,121 (SD = $480) in 2023.

Supplemental Table 3 presents a correlation matrix among variables of interest for the longitudinal sub-sample. Pandemic relief funding was correlated with itself over time (*r*'s range 0.39 – 0.51, *p*'s < 0.001), and greater funding was also associated with more child economic disadvantage across all years (*r*'s range 0.14 – 0.17, *p*'s < 0.001). A district that reported greater in-person learning in SY 2021 was also more likely to be rural ($r = 0.24$, $p < 0.001$), have smaller student memberships ($r = -0.33$, $p < 0.001$), and lower per pupil expenditures ($r = -0.18$, $p < 0.001$). These districts also received less pandemic relief funding in 2023 ($r = -0.19$, $p < 0.001$).

Cross-sectional Analyses

Achievement Recovery

Figure 2 displays average achievement scores for all students state-wide across the pre-pandemic and post-pandemic periods by cohort. Column 1 of Tables 2A and 2B displays these differences in reading achievement in raw units (Column 1 of Supplemental Table 4 reports these effects for math due to our lower confidence in the pre/post-pandemic math comparisons). Across all cohorts, students declined about 0.25 SDs in 2021 reading scores compared to their pre-pandemic benchmark, creating an achievement 'gap' between pre- versus post-pandemic test scores. By 2023, students remained ~0.15 SDs behind in reading achievement, demonstrating

some amount of recovery in reading scores. The immediate decline in math scores was much greater: across all cohorts, students averaged ~0.60 SDs lower on math achievement in 2021, compared to 2017 benchmarks. However, there was also evidence of greater recovery in math scores: by 2023, students were ~0.15 SDs behind 2019 benchmarks across all cohorts.

Developmental Timing of Exposure

Next, we examined differential patterns of recovery by student age. This was done by examining how much of the pandemic achievement 'gap' – i.e., the difference in test scores in 2021 compared to pre-pandemic benchmarks– was reduced by the end of 2023 in each cohort. Overall, we observed some evidence for differential patterns of recovery based on student age. The youngest cohort of students (i.e., Cohort 1; G3 in 2021) demonstrated the largest rates of recovery, closing 58% of the reading and 87% of the math achievement gap by G5. By 2023, Cohort 2 (G4 in 2021) closed 57% of reading and 67% of the math gap and Cohort 3 (G5 in 2021) closed 7% of reading and 63% of the math gap. Cohort 4 (G6 in 2021) closed 32% of the reading gap by 2023 (we cannot assess Cohort 4's math achievement as testing administration changed in grade 8).

Moderation by Economic Disadvantage

As anticipated, low-income students performed worse in reading (Column 2 of Tables 2A and 2B) and math (Column 2 of Supplemental Table 4) compared to their higher income peers after the pandemic. However, when examining patterns of moderation, we generally found positive interactions between the post-pandemic period indicator and the low-income indicator (Column 3 of Tables 2A and 2B). Careful interpretation of this interaction term is warranted. Both the low-income indicator and the post-COVID indicators produced large negative main effects in every model. This indicates that, on average, children from low-income families lagged

substantially behind their higher-income peers in both the pre- and post-COVID periods, and all children in the post-COVID period lagged significantly behind their pre-COVID peers. However, the positive interaction term indicates that lower-income students experienced a slightly reduced drop in 2021 achievement compared with higher-income students, and they also experienced slightly better recovery rates in 2022 and 2023. In standardized units, we observed, on average, that low-income students performed ~0.04 SDs better on reading achievement than higher-income peers' trends after the pandemic.

In contrast, low-income students fared worse on post-pandemic math achievement compared to their higher income peers across all cohorts (a reminder that we have lower confidence in these pre/post-pandemic math comparisons due to the test score change in 2019). As Column 3 of Supplemental Table 4 shows, on average, low-income students dropped ~0.08 SDs more on math achievement after the pandemic compared to higher income peers.

Longitudinal Analyses: COVID-era Policies and District-level Variation in Achievement

Table 3 presents variation in district-level reading achievement for a longitudinal cohort of students with both pre-pandemic and post-pandemic data. These models control for G3 test scores in 2019 to account for pre-pandemic levels of achievement. As the random intercept reflects, average reading scores across districts (after controlling for G3 scores) varied by about 25 points, and districts accounted for about 3% of the variance according to the ICC (Column 1). After adjusting for in-person learning, pandemic relief allocation, and student demographics, the variation in reading across districts declined modestly: about one-third (35%) of the variation in reading achievement in 2021 was explained once accounting for these factors, and the ICC declined to about 1% (Column 2). These factors accounted for slightly less district-level variation across later years, accounting for approximately 22% and 15% of reading achievement

variation in 2022 and 2023, respectively. Adjusting for other district characteristics did not explain additional variance in reading scores (Column 3).

In-person learning was associated with improved reading achievement, with each additional month spent in in-person instruction associated with a ~4-point increase on 2021 reading scores. This significant effect did not persist across 2022 and 2023. In contrast, greater pandemic relief funding was associated with poorer reading achievement: every $100 increase in pandemic relief funding per student was associated with ~1-point lower reading scores in 2021 and 2022. This association was not statistically significant in 2023.

Table 4 presents these findings for math achievement (of note, this analysis is not subject to the interpretative concerns of the cross-sectional math comparisons above as the same math exam was administered to students from 2019 through 2023). There was larger variation in math achievement than reading: math scores varied by about 45 points across districts, and districts explained about 8% of the variation according to the ICC (Column 1). Pandemic-era policies and student characteristics explained about one-third (30%) of this variation in 2021 and 15% of the variation in 2022 and 2023 (Column 2), with little additional variance explained via district demographics (Column 3). Broadly, in-person learning was strongly associated with improved math scores: each additional month spent in in-person instruction was associated with a ~10-point increase on G5 test scores in 2021. This effect did not broadly persist across additional years. Similar to reading scores, greater pandemic relief funding was associated with a small decline in math achievement.

Discussion

The present study leveraged population-level data from North Carolina public schools to describe pandemic-related recovery rates across middle childhood. Overall, we found greater

declines in students' math skills over reading achievement immediately following school reopening. There was strong evidence of recovery over time, though students did not reach pre-pandemic levels by the end of 2023. Recovery appeared stronger for math over reading achievement, which aligns with previous work documenting disparate impacts on children's math performance, but also greater rebounding of math scores over time compared to reading (Kuhfeld et al., 2022). We have less confidence in the precision of our math estimates due to a mean shift in math scores after an updated End-of-Grade exam was introduced in 2019; nonetheless, the magnitude of this shift and the alignment of our findings with the prior literature give us confidence in the general conclusion that math achievement was more responsive to the COVID-19 period than reading achievement.

Importantly, recovery was not universal across all students: younger students demonstrated the strongest rates of recovery, and low SES students demonstrated diverging patterns of recovery by subject matter. Finally, in-person instruction and pandemic relief funding explained modest variation in district-level achievement, suggesting the possible effects of these policies on recovery across the post-pandemic period.

Moderation by Economic Disadvantage

When examining moderation by student economic disadvantage, there were some surprising trends in low-income students' patterns of recovery. Counter to expectations, low-income students demonstrated slightly better recovery in reading compared to higher-income peers in all student cohorts; however, they also recovered worse in math (of note, we have lower confidence in the cross-sectional math results as there were noticeable differences in average math scores across all grades after the End-of-Grade test was updated in 2019, making pre- vs. post-pandemic achievement comparisons tenuous). This finding conflicts with previous work

which reported greater declines and decelerated growth rates for a similar population in both math and reading (Kuhfeld, Soland, Lewis, et al., 2022). Importantly, fewer economically disadvantaged students completed the End-of-Grade assessment across the post-pandemic period, which may bias results if lower-achieving students in these categories were more likely to attrit from public schools. In our investigations, we did not observe differences in attrition patterns by pre-pandemic achievement, though it is still possible that an achievement bias in attrition exists. Previous research also conducted in North Carolina found larger increases in chronic absenteeism for those in high-poverty schools after the pandemic (Swiderski et al., 2025), which may partially account for these students' missingness from the sample.

Developmental Timing of Exposure

There was substantial heterogeneity in recovery over time by student age. In line with our hypothesis, the youngest cohort of students (i.e., Grade 3 in 2021) demonstrated the strongest rate of recovery by 2023, closing over half of the reading gap and nearly 90% of the math gap. In contrast, the oldest cohort of students where both subject data were available closed less than 10% of the reading gap and about 70% of the math gap. These findings highlight the potential importance of child age in understanding the effects of environmental shocks to children's academic recovery. This aligns with the developmental theory that the repetitive nature of early instruction may have provided young children with the opportunity to "catch-up" at a faster rate than their older peers.

A meta-analysis of 39 studies found no statistically significant difference in the effect of the pandemic by student age (Di Pietro, 2023). However, this meta-analysis collapsed all primary school grades into a single category, whereas our data are able to examine heterogeneity *within* primary school. One way to interpret these findings is that elementary aged students may have

benefitted the most from the return to in-person learning. This aligns with recent work which found younger students demonstrated greater positive benefit from in-person instruction during lockdown (Jack et al., 2023).

Variation in Achievement by Pandemic-Era Policies

A notable contribution of our study is the examination of achievement variation across school districts. We observed large test score variation, which aligns with previous work: a case study of three metropolitan school districts observed similar variability in student achievement following school reopening (Sass & Goldring, 2021), and an examination of school-level variation found greater variability in math and reading scores after the onset of the pandemic within schools (Rambo-Hernandez et al., 2024). However, we extended these analyses further: we observed that the pandemic-era policies of in-person learning and pandemic relief funding explained about one-third of the district-level variation in test scores in SY 2021. This finding indicates that district-level decision-making on these policies may have influenced achievement outcomes immediately following school reopening. Indeed, previous work in North Carolina found that districts accounted for 43% of the variance in third graders' declining rate of math proficiency from 2019 to 2021, and increased achievement variance within schools (Domina et al., 2022), highlighting the potential importance of local decision making in student outcomes.

We observed that greater time in in-person learning was associated with increased achievement, with particularly strong associations with math scores in 2021. This aligns with descriptive work by Jack and colleagues (2023), which found that students in districts that offered greater access to in-person instruction had better pass rates on standardized exams. It is notable that these associations did not persist beyond 2021, highlighting the positive short-term

effects but relative fadeout of educational interventions that is pervasive in the literature (e.g., Hart et al., 2024).

In contrast, greater pandemic relief funding was associated with *lower* achievement in 2021. There may be several explanations to account for this association. On one hand, it is possible that this reflects a true causal relationship: districts that received greater funding perhaps did not spend funds appropriately, ultimately harming student achievement. However, this would be at odds with the recent literature on school funding or the theorized mechanisms by which funding impacts achievement (e.g., Jackson et al., 2015). Indeed, several studies of national data found small but positive impacts of pandemic relief aid on academic achievement (Dewey et al., 2025; Goldhaber & Falken, 2025). Importantly, greater dollars were allocated towards higher poverty districts, and these same districts demonstrated the greatest benefit of additional funding (Dewey et al., 2025; Goldhaber & Falken, 2025). Therefore, our observed negative association may be due to pandemic relief dollars being targeted towards more disadvantaged districts, and it's possible that our pre-pandemic test scores did not fully account for this selected disadvantage.

While we cannot derive strong causal inferences from these associations due to the possibility of selection bias, there are several important take-aways from this analysis. First, it is notable that district level policies are associated with decreased variance in achievement, which may indicate the importance of local decision making on student's short-term outcomes. In addition, we restricted this analysis to a longitudinal sub-sample where we could adjust for pre-pandemic achievement, removing a substantial amount of confounding variation. While this does not fully remove selection bias, it strengthens causal inference by answering the hypothetical

question ‘what is the effect of this policy on achievement for two children with the same baseline level of achievement before the pandemic?’

Strengths & Limitations

A strength of this work is the combination of state-wide cross-sectional analyses and within-child longitudinal analyses to describe pandemic recovery across middle childhood. We contribute to the extensive COVID-19 literature with the use of child-level data, pre-pandemic observations, and a novel test assessment specific to the state of North Carolina. In addition, we approach this analysis through a developmental perspective by considering child age as an important modifier of exogenous shocks on learning environments.

There are several limitations to this study. First, as a descriptive, observational study, we cannot derive causal inferences from the reported associations, particularly in regard to district policies. Second, due to the change in test form in SY 2019, we cannot definitely claim that scores across pre- and post-pandemic periods are completely comparable. While the Lexile and Quantile scales are designed to measure achievement above and beyond small variations in test items, this assumption did not appear to hold for Quantile scores (though we cannot fully rule out these measurement error concerns with the Lexile scale, it does appear more reliable due to the stability in scores across test iterations). Similarly, there were non-random demographic changes to schools’ population in the years after the pandemic. While we did not observe different likelihood of attrition in low-income children based on achievement, it is still possible that these estimates are biased. Third, there is a potential for measurement bias in the instructional mode variable, as this index was derived from monthly school-level data rather than individual students’ experiences. Finally, we cannot account for important family-level factors, such as household internet access, parental support, and student mental health, and district-level factors

such as how pandemic relief dollars were spent – all of which may have influenced learning outcomes during the pandemic.

Conclusion and Future Directions

The present study describes the current state of affairs in a large and diverse state in the United States. We observed that children in North Carolina are making continued progress toward recovery in reading and math achievement. Critically, school districts made key policy decisions during the COVID-19 pandemic in response to the needs of their communities. Nonetheless, these policy choices may have influenced the extent by which academic achievement was impacted during this time of crisis. Future research is necessary to identify which policies, if any, were most successful at buffering the negative effects of the COVID-19 pandemic on achievement, and how their impacts may vary across developmental stages. Identifying policies that mitigate the effect of such disruptive events on test scores and improve the rate of recovery are critical to informing the response to future crises.

Based on the findings of this study and others, it is clear that the COVID achievement gap remains a persistent and prescient challenge for schools. Moving forward, educators and policy makers must invest in best practices for improving academic skills to get students back on track. Some potentially promising areas for investment may include addressing chronic absenteeism, which has been associated with poorer recovery rates in several studies (Dee, 2024; Dewey et al., 2025). A recent research brief from Kuhfeld and Lewis (2025) encouraged states to consider investing in evidenced-based interventions, such as tutoring and summer programming, to boost test scores. Now that pandemic-era funding has ended, local educational leaders and policy makers play a vital role in determining the next steps of pandemic recovery and students' academic trajectories.

References

5 years after COVID-19 hit: Test data converge on math gains, stalled reading recovery. (n.d.). *Brookings*. Retrieved October 22, 2025, from https://www.brookings.edu/articles/5-years-after-covid-19-hit-test-data-converge-on-math-gains-stalled-reading-recovery/

Blackwell, C. K., Mansolf, M., Deoni, S. C. L., Ganiban, J. M., Leve, L. D., Margolis, A. E., McGrath, M., Nozadi, S. S., O'Shea, T. M., Sherlock, P., Zhao, Q., LeWinn, K. Z., & on behalf of program collaborators for Environmental influences on Child Health Outcomes Program. (2024). The impact of COVID-19 school disruptions on children's learning. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1295910

Cameron, C. E., Grimm, K. J., Steele, J. S., Castro-Schilo, L., & Grissmer, D. W. (2015). Nonlinear Gompertz curve models of achievement gaps in mathematics and reading. *Journal of Educational Psychology*, *107*(3), 789–804. https://doi.org/10.1037/edu0000009

Christian, A., Jacob, B., & Singleton, J. D. (2022). *Assessing School District Decision-Making: Evidence from the COVID-19 Pandemic* (Working Paper No. 30520). National Bureau of Economic Research. https://doi.org/10.3386/w30520

COVID Funds | NC DPI. (n.d.). Retrieved July 2, 2025, from https://www.dpi.nc.gov/districts-schools/district-operations/financial-and-business-services/covid-funds#Tab-Other-2709

COVID-19 School Data Hub. (n.d.). Retrieved July 16, 2025, from https://www.covidschooldatahub.com/

COVID-19 School Data Hub Team. (2021). *COVID-19 School Data Hub: Introduction and Data Availability Overview* [White Paper]. https://assets.ctfassets.net/9fbw4onh0qc1/51GumHMBAalCkUBBHREfxO/0f4e84f44f90451b5558e7989e0d342b/CSDH_Introductory_White_Paper.pdf

Curriculum Associates. (2023). *State of Student Learning in 2023* (Reading and Mathematics Annual Report). chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://cdn.bfldr.com/LS6J0F7/at/x8v8wp2c6j4s4wttsw2nwphb/ca-state-of-student-learning-technical-report-2023.pdf

Dee, T. S. (2024). Higher chronic absenteeism threatens academic recovery from the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *121*(3), e2312249121. https://doi.org/10.1073/pnas.2312249121

Dewey, D. C., Fahle, E. M., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2025). *Pivoting from Pandemic Recovery to Long-Term Reform* (Education Recovery Scorecard). chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://educationrecoveryscorecard.org/wp-content/uploads/2025/02/Pivoting-from-Pandemic-Recovery-to-Long-Term-Reform-A-District-Level-Analysis.pdf

Di Pietro, G. (2023). The impact of Covid-19 on student achievement: Evidence from a recent meta-analysis. *Educational Research Review*, *39*, 100530. https://doi.org/10.1016/j.edurev.2023.100530

Domina, T., Hashim, A., Kearney, C., Pham, L., & Smith, C. (2022). *COVID-19 and the System Resilience of Public Education: A View from North Carolina. An Essay for the Learning Curve*. Urban Institute. https://eric.ed.gov/?id=ED620779

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching students what they already know? The (mis)alignment between mathematics instructional content and student knowledge in kindergarten. *Educational Evaluation and Policy Analysis*, *35*(2), 157–178. https://doi.org/10.3102/0162373712461850

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17), e2022376118. https://doi.org/10.1073/pnas.2022376118

Fahle, E. M., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2024). *The First Year of Pandemic Recovery: A District-Level Analysis* (Education Recovery Scorecard). Center for Education Policy Research | The Educational Opportunity Project. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://educationrecoveryscorecard.org/wp-content/uploads/2024/01/ERS-Report-Final-1.31.pdf

Francis, D. J., Kulesz, P. A., & Benoit, J. S. (2018). Extending the Simple View of Reading to Account for Variation Within Readers and Across Texts: The Complete View of Reading (CVRi). *Remedial and Special Education*, *39*(5), 274–288. https://doi.org/10.1177/0741932518772904

Gajderowicz, T., Jakubowski, M., Kennedy, A., Christrup, C., Patrinos, H., & Strietholt, R. (2024). *The Learning Crisis: Three Years after Covid-19* (SSRN Scholarly Paper No. 5069978). Social Science Research Network. https://papers.ssrn.com/abstract=5069978

Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2023). The Educational Consequences of Remote and Hybrid Instruction during the Pandemic. *American Economic Review: Insights*, *5*(3), 377–392. https://doi.org/10.1257/aeri.20220180

Dan Goldhaber, Grace Falken (2025). ESSER and Student Achievement: Assessing the Impacts of the Largest One-Time Federal Investment in K12 Schools. CALDER Working Paper No. 301-0325-2

Halloran, C., Jack, R., Okun, J. C., & Oster, E. (2021). *Pandemic Schooling Mode and Student Test Scores: Evidence from US States* (Working Paper No. 29497). National Bureau of Economic Research. https://doi.org/10.3386/w29497

Hart, E. R., Bailey, D. H., Luo, S., Sengupta, P., & Watts, T. W. (2024). Fadeout and persistence of intervention impacts on social–emotional and cognitive skills in children and adolescents: A meta-analytic review of randomized controlled trials. *Psychological Bulletin*, *150*(10), 1207–1236. https://doi.org/10.1037/bul0000450

Jack, R., Halloran, C., Okun, J., & Oster, E. (2023). Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts. *American Economic Review: Insights*, *5*(2), 173–190. (2041093). https://doi.org/10.1257/aeri.20210748

Jackson, C. K., Johnson, R. C., & Persico, C. (2015). *The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms* (Working Paper No. 20847). National Bureau of Economic Research. https://doi.org/10.3386/w20847

Kuhfeld, M., & Lewis, K. (2022). *Student achievement in 2021–2022: Cause for hope and continued urgency*. NWEA Research.

Kuhfeld, M., Lewis, K., & Peltier, T. (2023). Reading achievement declines during the COVID-19 pandemic: Evidence from 5 million U.S. students in grades 3–8. *Reading and Writing*, *36*(2), 245–261. https://doi.org/10.1007/s11145-022-10345-8

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test Score Patterns Across Three COVID-19-Impacted School Years. *Educational Researcher*, *51*(7), 500–506. https://doi.org/10.3102/0013189X221109178

Kuhfeld, M., Soland, J., Lewis, K., Ruzek, E., & Johnson, A. (2022). The COVID-19 School Year: Learning and Recovery Across 2020-2021. *AERA Open*, *8*, 23328584221099306. https://doi.org/10.1177/23328584221099306

Lewis, K., & Kuhfeld, M. (2023). *Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery*.

Mok, M. M. C., McInerney, D. M., Zhu, J., & Or, A. (2015). Growth trajectories of mathematics achievement: Longitudinal tracking of student academic progress. *British Journal of Educational Psychology*, *85*(2), 154–171. https://doi.org/10.1111/bjep.12060

National Center for Education Statistics. (n.d.). *Common Core of Data* [Dataset]. Retrieved July 24, 2025, from https://nces.ed.gov/ccd/files.asp#Fiscal:1,LevelId:5,SchoolYearId:37,Page:1

Osborne, M., Clifford, B., Faulk, P., Gilmore, D., & Smith, C. (2020). *NC Districts Fall 2020 Reopening Database* [Dataset]. EducationNC & Public Impact. https://docs.google.com/spreadsheets/d/1A2VKl0Y3OVPF42vRaA54xItnKlYAzXN22eUrfbKWMOA/edit?gid=176684338#gid=176684338

Osborne, M., Rash, M., Bryan, A., Clifford, B., Dejohn, M., Halstater, K., Struhs, J., Whalen, J., & Smith, T. (2020). *NC District Responses to COVID-19 School Closures* [Dataset]. EducationNC & Public Impact.

https://docs.google.com/spreadsheets/d/1eAwTyU5wbltxbvNaxkOl2bmTg0WcyfZ0S40RAgBw3UU/edit?gid=1202431252#gid=1202431252

Peters, S. J., Langi, M., Kuhfeld, M., & Lewis, K. (2025). Unequal Learning Loss: How the COVID-19 Pandemic Influenced the Academic Growth of Learners at the Tails of the Achievement Distribution. *Journal for the Education of the Gifted*, *48*(1), 25–47. https://doi.org/10.1177/01623532241301079

Rambo-Hernandez, K. E., Makel, M. C., & Koehler, N. (2024). Millions of Students Are (Still) Above Grade Level: Achievement and Achievement Variability in Mathematics and Reading Before and During COVID-19 in the United States. *Journal for the Education of the Gifted*, *47*(4), 385–409. https://doi.org/10.1177/01623532241277840

Sanford-Moore, E. (2022). *Assessment Features That Make a Difference When Examining Growth* (Features of Assessments and Interim-Summative Growth) [Position Paper]. MetaMetrics. https://metametricsinc.com/wp-content/uploads/2022/04/Assessment-Features-That-Make-a-Difference-When-Examining-Growth.pdf

Sass, T., & Ali, S. M. (2024, March 4). Student Achievement Growth During the COVID-19 Pandemic: Spring 2023 Update. *Georgia Policy Labs*. https://gpl.gsu.edu/publications/student-achievement-growth-during-the-covid-19-pandemic-spring-2023-update/

Sass, T., & Goldring, T. (2021). Student Achievement Growth During the COVID-19 Pandemic. *Georgia Policy Labs*. https://doi.org/10.57709/30728958

Swiderski, T., Fuller, S. C., & Bastian, K. C. (2025). Student-Level Attendance Patterns Across Three Post-Pandemic Years. *Educational Evaluation and Policy Analysis*, 01623737251315715. https://doi.org/10.3102/01623737251315715

Troppe, P., Isenberg, E., Rizzo, L., Standing, K., Deacon, G., Ross, C., Gill, B. P., Li, A., & Williams, B. (2024). *Implementation of Key Federal Education Policies in the Wake of the Coronavirus Pandemic: State and District Actions During the 2020–21 School Year (NCEE 2025-003c)*. U.S.

Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee

Table 1. Sample descriptives across pre- and post-pandemic cohorts of students

	Pre-Pandemic (SY 2017-2019)	Post-Pandemic (SY 2021-2023)	Point Change	p-value
Cohort 1 – Grades 3, 4, 5				
Student is Male	51.04	51.29	0.25	0.079
Whether non-Hispanic Black	24.65	22.93	-1.71	<0.001
Whether Hispanic	19.75	21.78	2.03	<0.001
Whether Other	8.82	10.35	1.53	<0.001
Whether non-Hispanic White	46.79	44.94	-1.85	<0.001
Special education status	12.88	13.10	0.22	0.019
Ever repeated a grade	0.88	0.35	-0.53	<0.001
English language learner	13.77	13.83	0.06	0.571
Economically disadvantaged	50.12	45.31	-4.81	<0.001
N	279,262	222,876		
Cohort 2 – Grades 4, 5, 6				
Student is Male	51.18	51.10	-0.08	0.570
Whether non-Hispanic Black	24.26	23.28	-0.98	<0.001
Whether Hispanic	20.17	21.71	1.55	<0.001
Whether Other	8.79	10.39	1.60	<0.001
Whether non-Hispanic White	46.78	44.62	-2.16	<0.001
Special education status	12.89	12.49	-0.40	<0.001
Ever repeated a grade	1.51	0.83	-0.68	<0.001
English language learner	13.65	14.17	0.52	<0.001
Economically disadvantaged	49.95	45.22	-4.73	<0.001
N	274,524	217,944		
Cohort 3 – Grades 5, 6, 7				
Student is Male	50.97	51.37	0.40	0.005
Whether non-Hispanic Black	24.13	23.43	-0.70	<0.001
Whether Hispanic	19.57	22.13	2.55	<0.001
Whether Other	8.67	10.11	1.44	<0.001
Whether non-Hispanic White	47.63	44.34	-3.29	<0.001
Special education status	13.02	12.25	-0.77	<0.001
Ever repeated a grade	1.95	1.07	-0.88	<0.001
English language learner	13.53	15.12	1.59	<0.001
Economically disadvantaged	49.06	44.40	-4.66	<0.001
N	268,293	222,960		
Cohort 4 – Grades 6, 7, 8				
Student is Male	50.91	51.48	0.57	<0.001
Whether non-Hispanic Black	23.87	23.79	-0.08	0.508
Whether Hispanic	18.48	21.95	3.46	<0.001
Whether Other	8.42	9.82	1.40	<0.001
Whether non-Hispanic White	49.23	44.44	-4.79	<0.001
Special education status	12.18	11.60	-0.58	<0.001
Ever repeated a grade	1.88	1.33	-0.56	<0.001
English language learner	13.20	15.57	2.37	<0.001
Economically disadvantaged	46.86	43.57	-3.28	<0.001
N	268,644	235,723		

Note. Sample includes all students with 3 repeated observations of reading and math test scores within the same school district. P-value reflects the results from a t-test examining the mean difference in student demographic characteristics in the pre/post-pandemic periods. "Other" category reflects students identifying as Asian, American Indian/Alaskan Native, or Native Hawaiian/Pacific Islander. Economically disadvantaged reflects whether a student is eligible for free/reduced price lunch.

Figure 1. State-wide achievement score trends by grade
Figure 1A: Reading

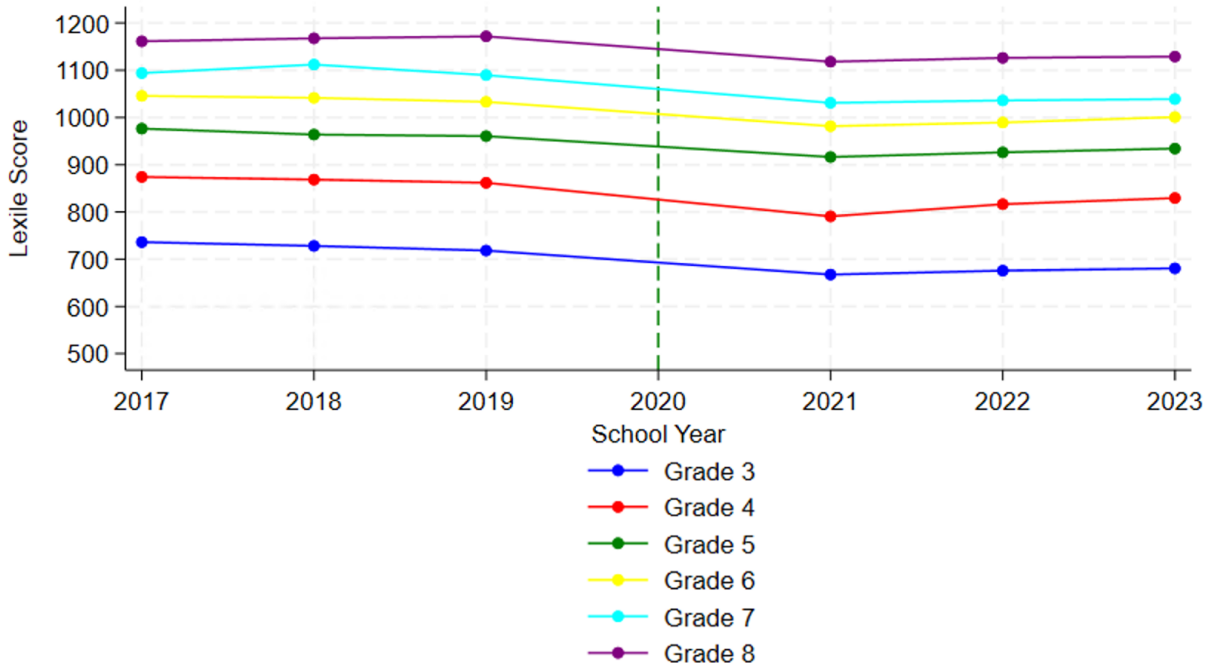

Figure 1B: Math

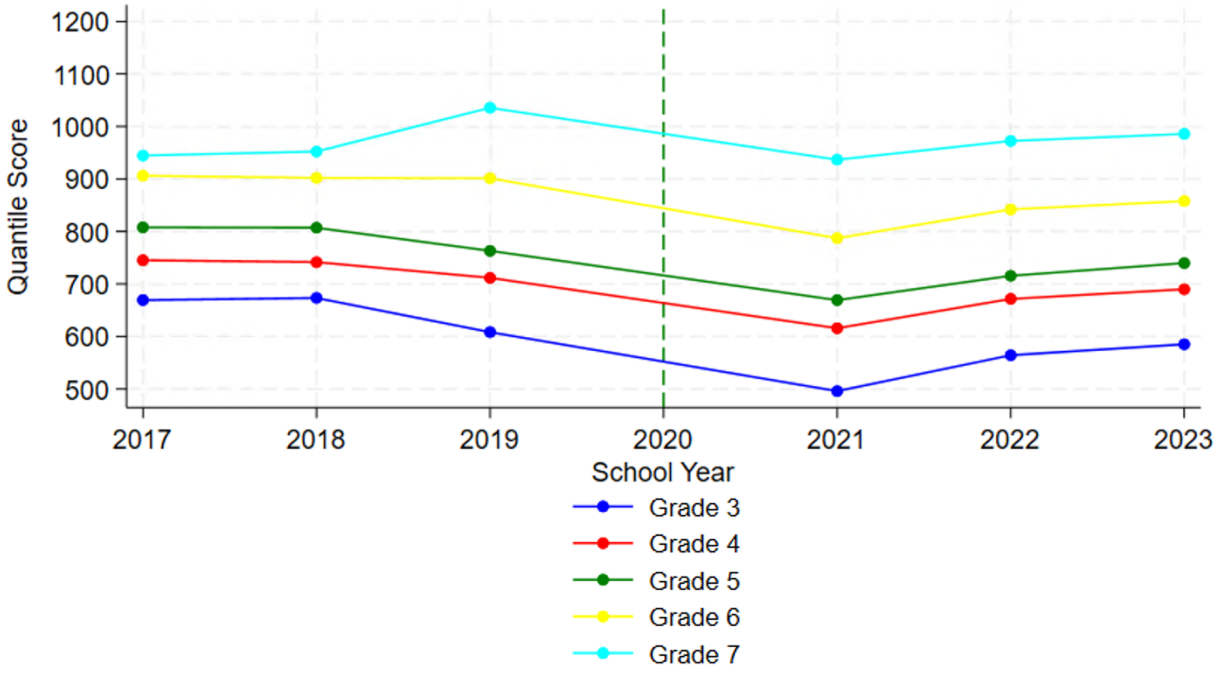

Figure 2. Growth in achievement across pre-pandemic and post-pandemic cohorts in raw units
Figure 2A. Lexile Scores

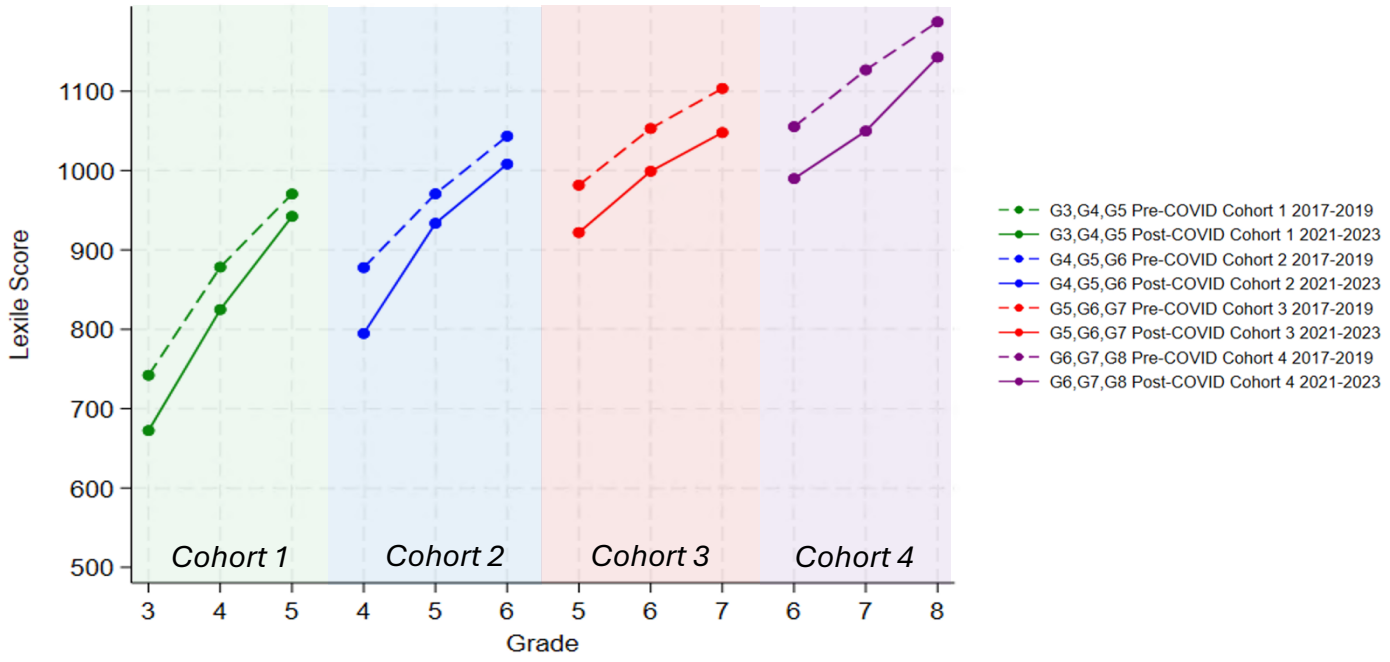

Figure 2B. Quantile Scores

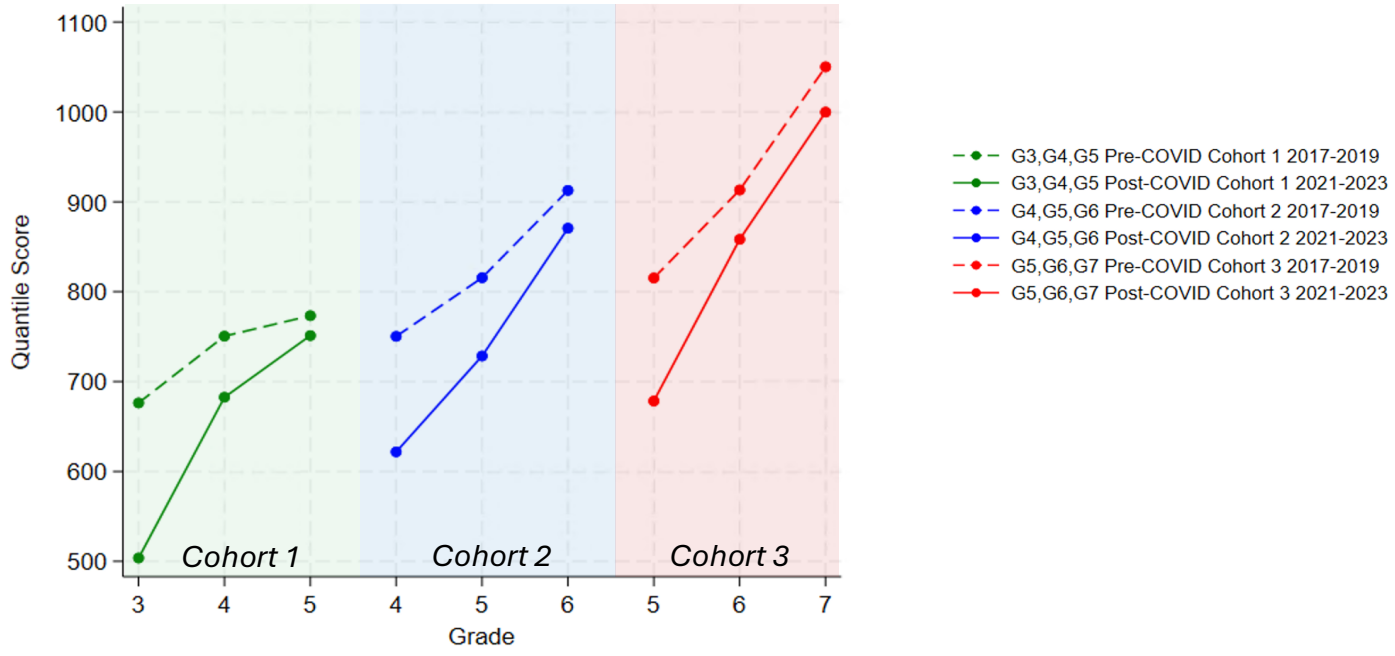

Table 2A. Reading achievement across pre- and post-pandemic periods (Cohorts 1 and 2)

	Statewide Reading Achievement: Cohort 1 (G3, G4, G5)		
N=167,373	(1)	(2)	(3)
Grade 3 (2017 vs 2021)			
Post-COVID	-70.70*** (1.20)	-85.95*** (1.03)	-88.02*** (1.41)
Low SES		-100.10*** (1.13)	-102.09*** (1.46)
Post-COVID X Low SES			4.41* (2.06)
Grade 4 (2018 vs 2022)			
Post-COVID	-54.30*** (1.18)	-57.23*** (0.99)	-59.96*** (1.32)
Low SES		-93.25*** (1.07)	-96.00*** (1.38)
Post-COVID X Low SES			6.26** (1.98)
Grade 5 (2019 vs 2023)			
Post-COVID	-29.68*** (1.15)	-23.79*** (0.96)	-23.85*** (1.38)
Low SES		-95.62*** (1.06)	-95.66*** (1.36)
Post-COVID X Low SES			0.11 (1.91)
	Statewide Reading Achievement: Cohort 2 (G4, G5, G6)		
N=164,156	(1)	(2)	(3)
Grade 4 (2017 vs 2021)			
Post-COVID	-83.86*** (1.18)	-98.73*** (1.01)	-100.63*** (1.37)
Low SES		-94.20*** (1.10)	-96.04*** (1.42)
Post-COVID X Low SES			4.11* (2.00)
Grade 5 (2018 vs 2022)			
Post-COVID	-38.13*** (1.15)	-42.00*** (0.96)	-43.90*** (1.27)
Low SES		-86.77*** (1.04)	-88.71*** (1.34)
Post-COVID X Low SES			4.42* (1.93)
Grade 6 (2019 vs 2023)			
Post-COVID	-36.01*** (1.25)	-31.42*** (1.05)	-37.04*** (1.52)
Low SES		-107.69*** (1.17)	-112.49*** (1.50)
Post-COVID X Low SES			10.74*** (2.09)
Includes Student-level Demographics?		X	X
Includes District-level Demographics?			X

Note. *p<0.05, **p<0.01, ***p<0.001. Mixed models estimating average achievement in post-pandemic years compared to pre-pandemic reference points. Models are nested within school districts. Column 1 presents the average achievement decline in each post-pandemic year. Column 2 introduces adjustments for the following student demographics: child sex, race/ethnicity, eligibility for free/reduced price lunch, special education status, English language learner status, and whether the child has ever repeated a grade. Column 3 examines moderation of effects via a Post-COVID X Low-SES interaction term and further adjusts for the following district characteristics: whether rural, student-teacher ratio, district membership, and total per pupil expenditure. Sample includes all students with a three valid test scores (i.e., observed in each grade) across either the pre-pandemic (2017-2019) or post-pandemic (2021-2023) period (see Table 1).

Table 2B. Reading achievement across pre- and post-pandemic periods (Cohorts 3 and 4)

	Statewide Reading Achievement: Cohort 3 (G5, G6, G7)		
N=163,751	(1)	(2)	(3)
Grade 5 (2017 vs 2021)			
Post-COVID	-60.33*** (1.14)	-72.88*** (0.97)	-74.30*** (1.31)
Low SES		-89.87*** (1.06)	-91.29*** (1.38)
Post-COVID X Low SES			3.11 (1.93)
Grade 6 (2018 vs 2022)			
Post-COVID	-54.18*** (1.24)	-59.93*** (1.05)	-69.75*** (1.38)
Low SES		-96.38*** (1.14)	-106.65*** (1.48)
Post-COVID X Low SES			22.78*** (2.10)
Grade 7 (2019 vs 2023)			
Post-COVID	-56.35*** (1.28)	-50.21*** (1.08)	-56.00*** (1.53)
Low SES		-105.02*** (1.20)	-110.24*** (1.55)
Post-COVID X Low SES			11.38*** (2.14)

	Statewide Reading Achievement: Cohort 4 (G6, G7, G8)		
N=168,119	(1)	(2)	(3)
Grade 6 (2017 vs 2021)			
Post-COVID	-65.06*** (1.20)	-74.96*** (1.03)	-86.93*** (1.39)
Low SES		-97.65*** (1.13)	-110.19*** (1.49)
Post-COVID X Low SES			26.43*** (2.05)
Grade 7 (2018 vs 2022)			
Post-COVID	-76.83*** (1.25)	-78.76*** (1.06)	-85.42*** (1.38)
Low SES		-95.22*** (1.17)	-102.69*** (1.53)
Post-COVID X Low SES			16.12*** (2.14)
Grade 8 (2019 vs 2023)			
Post-COVID	-44.30*** (1.25)	-34.49*** (1.06)	-38.24*** (1.48)
Low SES		-102.16*** (1.18)	-105.82*** (1.55)
Post-COVID X Low SES			7.69*** (2.11)
Includes Student-level Demographics?		X	X
Includes District-level Demographics?			X

Note. *p<0.05, **p<0.01, ***p<0.001. Mixed models estimating average achievement in post-pandemic years compared to pre-pandemic reference points. Models are nested within school districts. Column 1 presents the average achievement decline in each post-pandemic year. Column 2 introduces adjustments for the following student demographics: child sex, race/ethnicity, eligibility for free/reduced price lunch, special education status, English language learner status, and whether the child has ever repeated a grade. Column 3 examines moderation of effects via a Post-COVID X Low-SES interaction term and further adjusts for the following district characteristics: whether rural, student-teacher ratio, district membership, and total per pupil expenditure. Sample includes all students with a three valid test scores (i.e., observed in each grade) across either the pre-pandemic (2017-2019) or post-pandemic (2021-2023) period (see Table 1).

Table 3. Association between pandemic policies and within-child reading achievement

	Within-Child Reading Achievement		
N=73,289	(1)	(2)	(3)
Grade 5 (2021)			
In-person Learning (2021)		3.43** (1.26)	4.04** (1.37)
Pandemic Relief Dollars (2021)		-1.49** (0.52)	-1.32* (0.56)
District ICC	0.03	0.01	0.01
District Random Intercept SD	25.13	16.43	15.93
Grade 6 (2022)			
In-person Learning (2021)		-1.31 (1.39)	-0.23 (1.50)
Pandemic Relief Dollars (2022)		-0.86** (0.31)	-0.85** (0.32)
District ICC	0.02	0.01	0.01
District Random Intercept SD	23.28	18.17	17.34
Grade 7 (2023)			
In-person Learning (2021)		-0.03 (1.66)	0.91 (1.78)
Pandemic Relief Dollars (2023)		0.22 (0.23)	0.27 (0.24)
District ICC	0.03	0.02	0.02
District Random Intercept SD	25.74	22.12	21.03
Includes Pre-COVID Achievement?	X	X	X
Includes Student-level Demographics?		X	X
Includes District-level Demographics?			X

Note. *p<0.05, **p<0.01, ***p<0.001. Mixed models estimating the association between district-level policies and student achievement after COVID-19. Models are nested within school districts. Column 1 reflects achievement regressed on a Post-COVID indicator for each grade within each cohort. Column 2 introduces associations with in-person learning and pandemic relief funding, with adjustments for the following student demographics: child sex, race/ethnicity, eligibility for free/reduced price lunch, special education status, English language learner status, and whether the child has ever repeated a grade. Column 3 further adjusts for the following district characteristics: whether rural, student-teacher ratio, district membership, and total per pupil expenditure. Models control for pre-pandemic (i.e., Grade 3) achievement in 2019. The In-Person Learning index is scaled in months, while Pandemic Relief Dollars are scaled in hundreds of dollars allocated per student per district. Analyses restricted to a longitudinal sub-sample of students with data available across pre-pandemic (Grade 3) and post-pandemic periods (Grades 5 through 7; see Supplemental Table 1).

Table 4. Association between pandemic policies and within-child math achievement

	Within-Child Math Achievement		
N=73,289	(1)	(2)	(3)
Grade 5 (2021)			
In-person Learning (2021)		9.92*** (2.07)	11.54*** (2.20)
Pandemic Relief Dollars (2021)		-2.10*** (0.62)	-1.84** (0.64)
District ICC	0.09	0.05	0.04
District Random Intercept SD	41.60	29.45	27.83
Grade 6 (2022)			
In-person Learning (2021)		2.98 (2.95)	5.03 (3.11)
Pandemic Relief Dollars (2022)		-0.69 (0.40)	-0.53 (0.40)
District ICC	0.08	0.07	0.06
District Random Intercept SD	50.04	42.72	39.85
Grade 7 (2023)			
In-person Learning (2021)		3.38 (2.91)	6.09* (3.03)
Pandemic Relief Dollars (2023)		-0.92** (0.28)	-0.84** (0.28)
District ICC	0.07	0.06	0.05
District Random Intercept SD	48.85	41.54	38.34
Includes Pre-COVID Achievement?	X	X	X
Includes Student-level Demographics?		X	X
Includes District-level Demographics?			X

Note. *p<0.05, **p<0.01, ***p<0.001. Mixed models estimating the association between district-level policies and student achievement after COVID-19. Models are nested within school districts. Column 1 reflects achievement regressed on a Post-COVID indicator for each grade within each cohort. Column 2 introduces associations with in-person learning and pandemic relief funding, with adjustments for the following student demographics: child sex, race/ethnicity, eligibility for free/reduced price lunch, special education status, English language learner status, and whether the child has ever repeated a grade. Column 3 further adjusts for the following district characteristics: whether rural, student-teacher ratio, district membership, and total per pupil expenditure. Models control for pre-pandemic (i.e., Grade 3) achievement in 2019. The In-Person Learning index is scaled in months, while Pandemic Relief Dollars are scaled in hundreds of dollars allocated per student per district. Analyses restricted to a longitudinal sub-sample of students with data available across pre-pandemic (Grade 3) and post-pandemic periods (Grades 5 through 7; see Supplemental Table 1).

Supplemental Materials

Supplemental Table 1. Descriptives of full sample and analytic sub-samples

	Full Sample	Cross-Sectional Sample	Longitudinal Sample
Student is Male	51.20%	51.15%	51.19%
Whether non-Hispanic Black	24.74%	23.84%	24.30%
Whether Hispanic	20.15%	20.74%	21.74%
Whether non-Hispanic Other	9.88%	9.44%	9.92%
Whether non-Hispanic White	45.23%	45.97%	44.04%
Special education status	12.90%	12.79%	12.17%
Ever repeated a grade	1.64%	1.13%	1.25%
English language learner	13.76%	13.97%	14.55%
Economically disadvantaged	47.89%	47.60%	47.54%
	3,101,929	1,990,226	293,156

Note. "Full sample" refers to all individual records observed between 2017 – 2019 and 2021 – 2023. "Cross-sectional sample" refers to all individual records with at least 3 repeated assessments of reading and math in either the pre-pandemic (2017-2019) or post-pandemic period (2021-2023) within the same public school district. "Longitudinal sample" refers to all individual records of students observed in both the pre-pandemic (i.e., G3 in 2019) and post-pandemic period (G5 in 2021, G6 in 2022, and G7 in 2023). P-value reflects the results from a t-test examining the mean difference in student demographic characteristics for each group.

Supplemental Table 2. Predicting odds of attrition from 2019 to 2021 based on student characteristics

	Odds Ratio	95% C.I.
Quantile Scores	1.00 ***	[1.00, 1.00]
Lexile Scores	0.99 ***	[0.99, 1.00]
Student is Male	1.02 *	[1.00, 1.04]
Whether non-Hispanic Black	1.07 ***	[1.05, 1.10]
Whether Hispanic	0.87 ***	[0.84, 0.89]
Whether non-Hispanic Other	1.01	[0.98, 1.05]
Special education status	1.04 **	[1.01, 1.07]
Ever repeated a grade	1.20 ***	[1.09, 1.26]
English language learner	0.89 ***	[0.86, 0.89]
Economically disadvantaged	0.88 ***	[0.86, 0.89]
N	329,561	

Note. *$p<0.05$, **$p<0.01$, ***$p<0.001$. Logistic regression estimating the likelihood of remaining in the sample from 2019 to 2021, where 0=remain enrolled and 1=attrit. Coefficients reflect odds ratios, where a value > 1 signifies a greater risk of attrition. Analysis restricted to students in Grades 3, 4, or 5 in SY 2019.

Supplemental Table 3: Correlations amongst district and student characteristics in the longitudinal subsample

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) In-person Index	1.00															
(2) Pandemic Funds 2021	0.02	1.00														
(3) Pandemic Funds 2022	-0.03	0.39	1.00													
(4) Pandemic Funds 2023	-0.19	0.51	0.43	1.00												
(5) District is Rural	0.24	0.19	0.12	0.02	1.00											
(6) Student-Teacher Ratio	0.07	-0.21	-0.42	-0.17	-0.18	1.00										
(7) District Membership	-0.33	-0.38	-0.22	-0.17	-0.46	0.41	1.00									
(8) Per Pupil Expenditures	-0.18	0.32	0.39	0.23	0.09	-0.74	-0.27	1.00								
(9) Child is Male	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00							
(10) Child is Black	-0.13	0.17	0.10	0.20	-0.08	0.02	0.06	0.08	-0.01	1.00						
(11) Child is Hispanic	-0.04	-0.01	-0.01	0.02	-0.02	0.02	0.05	-0.01	0.00	-0.30	1.00					
(12) Child is Other Race	-0.06	0.00	0.01	-0.02	-0.03	0.02	0.06	0.00	0.00	-0.19	-0.17	1.00				
(13) Special Ed Status	0.02	0.02	0.02	0.01	0.02	-0.03	-0.04	0.01	0.11	0.03	-0.02	-0.03	1.00			
(14) Child Repeated Grade	-0.01	0.04	0.01	0.04	0.02	-0.01	-0.02	0.03	0.01	0.04	0.00	-0.01	0.02	1.00		
(15) Limited English Proficiency	-0.05	-0.01	0.00	0.01	-0.04	0.02	0.08	0.00	0.02	-0.21	0.62	0.01	0.00	0.01	1.00	
(16) Economically Disadvantaged	0.01	0.17	0.14	0.14	0.10	-0.09	-0.16	0.07	0.00	0.25	0.17	-0.03	0.09	0.04	0.14	1.00

Note. N=73,289. Variables 1 – 8 reflect district-level characteristics. Variables 9 – 16 reflect child-level characteristics.

Supplemental Table 4: Math achievement across pre- and post-pandemic periods (Cohorts 1, 2, and 3)

	Statewide Math Achievement: Cohort 1 (G3, G4, G5)		
N=164,156	(1)	(2)	(3)
Grade 3 (2017 vs 2021)			
Post-COVID	-174.37*** (1.11)	-188.94*** (0.98)	-172.85*** (1.34)
Low SES		-100.50*** (1.07)	-85.02*** (1.39)
Post-COVID X Low SES			-34.35*** (1.95)
Grade 4 (2018 vs 2022)			
Post-COVID	-69.50*** (1.02)	-73.62*** (0.88)	-70.10*** (1.17)
Low SES		-87.24*** (0.95)	-83.69*** (1.23)
Post-COVID X Low SES			-8.07*** (1.76)
Grade 5 (2019 vs 2023)			
Post-COVID	-23.54*** (1.06)	-18.63*** (0.90)	-8.78*** (1.29)
Low SES		-96.53*** (0.99)	-87.97*** (1.27)
Post-COVID X Low SES			-19.10*** (1.78)
	Statewide Math Achievement: Cohort 2 (G4, G5, G6)		
N=167,385	(1)	(2)	(3)
Grade 4 (2017 vs 2021)			
Post-COVID	-129.87*** (1.01)	-143.03*** (0.88)	-139.78*** (1.20)
Low SES		-86.78*** (0.96)	-83.64*** (1.25)
Post-COVID X Low SES			-6.99*** (1.76)
Grade 5 (2018 vs 2022)			
Post-COVID	-88.46*** (1.06)	-93.45*** (0.91)	-89.77*** (1.20)
Low SES		-88.98*** (0.99)	-85.23*** (1.27)
Post-COVID X Low SES			-8.57*** (1.82)
Grade 6 (2019 vs 2023)			
Post-COVID	-43.34*** (1.32)	-38.58*** (1.12)	-27.56*** (1.61)
Low SES		-129.27*** (1.24)	-119.85*** (1.59)
Post-COVID X Low SES			-21.06*** (2.22)
	Statewide Math Achievement: Cohort 3 (G5, G6, G7)		
N=163,751	(1)	(2)	(3)
Grade 5 (2017 vs 2021)			
Post-COVID	-138.04*** (1.04)	-150.57*** (0.91)	-142.30*** (1.22)
Low SES		-86.75*** (0.99)	-78.42*** (1.29)
Post-COVID X Low SES			-18.15*** (1.81)
Grade 6 (2018 vs 2022)			
Post-COVID	-55.21*** (1.17)	-62.22*** (1.00)	-45.27*** (1.32)
Low SES		-99.57*** (1.10)	-81.83*** (1.42)
Post-COVID X Low SES			-39.31*** (2.01)
Grade 7 (2019 vs 2023)			
Post-COVID	-50.96*** (1.33)	-44.97*** (1.14)	-35.22*** (1.62)
Low SES		-126.29*** (1.26)	-117.48*** (1.63)
Post-COVID X Low SES			-19.17*** (2.26)
Includes Student-level Demographics?		X	X
Includes District-level Demographics?			X

Note. *p<0.05, **p<0.01, ***p<0.001. Mixed models estimating average achievement in post-pandemic years compared to pre-pandemic reference points. Models are nested within school districts. Column 1 presents the average achievement decline in each post-pandemic year. Column 2 introduces adjustments for the following student demographics: child sex, race/ethnicity, eligibility for free/reduced price lunch, special education status, English language learner status, and whether the child has ever repeated a grade. Column 3 examines moderation of effects via a Post-COVID X Low-SES interaction term. Sample includes all students with a three valid test scores (i.e., observed in each grade) across either the pre-pandemic (2017-2019) or post-pandemic (2021-2023) period.