ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 24-951

The relationship between student attendance and achievement, pre- and post-COVID

Tom Swiderski
University of North Carolina at Chapel Hill

Sarah Crittenden Fuller
University of North Carolina at Chapel Hill

Kevin C. Bastian
University of North Carolina at Chapel Hill

We examine the relationship between absenteeism and achievement since the onset of COVID-19. Applying first-differences models to North Carolina administrative data, we estimate that each absence was associated with a 0.0032 standard deviation (SD) decline in math achievement in 2022-23. As students averaged 3.3 more absences in 2022-23 than 2018-19, these results imply that returning absence rates to pre-pandemic levels in 2022-23 may have increased overall achievement by 0.010 SDs. We identify a stronger relationship between absenteeism and achievement pre-pandemic. Across three pre-pandemic cohorts, each absence was associated with a 0.0055 SD decline in achievement, with some evidence of additional impacts of peer absenteeism. Results from these models imply that returning attendance to pre-pandemic levels could have improved 2022-23 achievement by 0.018–0.031 SDs. Findings highlight the important but partial role that attendance recovery may play in academic recovery and suggest potential changes in the costliness of absenteeism post-pandemic.

VERSION: November 2024

Suggested citation: Swiderski, Tom, Sarah Crittenden Fuller, and Kevin C. Bastian. (2024). The relationship between student attendance and achievement, pre- and post-COVID. (EdWorkingPaper: 24 -951). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/t1dp-tn53

The relationship between student attendance and achievement, pre- and post-COVID

Tom Swiderski[1]
tswid@email.unc.edu

Sarah Crittenden Fuller[1]
Sarah.fuller@unc.edu

Kevin C. Bastian[1]
kbastian@email.unc.edu

[1] University of North Carolina at Chapel Hill

Abstract

We examine the relationship between absenteeism and achievement since the onset of COVID-19. Applying first-differences models to North Carolina administrative data, we estimate that each absence was associated with a 0.0032 standard deviation (SD) decline in math achievement in 2022-23. As students averaged 3.3 more absences in 2022-23 than 2018-19, these results imply that returning absence rates to pre-pandemic levels in 2022-23 may have increased overall achievement by 0.010 SDs. We identify a stronger relationship between absenteeism and achievement pre-pandemic. Across three pre-pandemic cohorts, each absence was associated with a 0.0055 SD decline in achievement, with some evidence of additional impacts of peer absenteeism. Results from these models imply that returning attendance to pre-pandemic levels could have improved 2022-23 achievement by 0.018–0.031 SDs. Findings highlight the important but partial role that attendance recovery may play in academic recovery and suggest potential changes in the costliness of absenteeism post-pandemic.

Keywords: COVID-19; attendance; achievement

Acknowledgements: We thank the North Carolina Department of Public Instruction (NCDPI) for providing data used in this work. This research was supported by the Institute of Education Sciences (IES), U.S. Department of Education, through grant number R305S220003 to NCDPI, as well as a grant from the Spencer Foundation (#202200138). The views expressed are those of the authors and do not necessarily reflect the views of NCDPI, IES, the U.S. Department of Education, or the Spencer Foundation.

I. Introduction

This study aims to estimate the impact of post-pandemic student absenteeism on achievement to determine the extent to which heightened absenteeism may be slowing the pace of K-12 academic recovery. This work is motivated by three existing findings. First, due to the COVID-19 pandemic, student test scores dropped markedly in 2020-21 and have since recovered only partially (Curriculum Associates, 2023; Fahle et al., 2024; Kuhfeld & Lewis, 2024; Lewis & Kuhfeld, 2023). Second, post-pandemic absenteeism is a growing concern. Between 2018-19 and 2021-22, the chronic absence rate doubled nationally from 15% to 28% (Dee, 2024) and has since declined only modestly (Malkus, 2024). Third, absenteeism has historically been shown to have modest but significant adverse effects on a student's own achievement and can further spill over to negatively impact the achievement of classroom peers (e.g., Aucejo & Romano, 2016; Gershenson et al., 2017; Gottfried & Ansari, 2022).

Together, these findings highlight the important role that attendance recovery may play in post-pandemic academic recovery. However, there is little work quantifying the impact of post-pandemic absenteeism on student achievement. One report found that rising absenteeism may explain 16% to 27% of the decline in math scores on the National Assessment of Educational Progress, or NAEP (National Center for Education Statistics [NCES], 2023). This analysis was based on students' self-reported attendance in the month prior to the exam and adjusts for changes in student demographic characteristics.

The current study extends this work by using student-level administrative data from North Carolina to pursue two main goals. First, we estimate the relationship between absenteeism and math achievement using student first-differences models in a cohort of elementary and middle school students observed between 2021-22 and 2022-23. We use this result to obtain a back-of-the-envelope estimate of how much better test scores in 2022-23 might have been if absenteeism had returned to pre-pandemic levels in this year. Second, we examine whether the relationship between absences and test scores has changed from pre- to post-pandemic. For example, the increasing availability of classroom resources online (Mervosh & Paris, 2024; Jacob & Stanojevich, 2024) may make each absence less costly than pre-pandemic. Therefore, we replicate our model in three pre-pandemic cohorts.

We find that each additional absence is associated with a 0.0032 standard deviation (SD) decline in math achievement in our post-pandemic cohort. As students averaged 3.27 additional absences in 2022-23 relative to pre-pandemic, results imply that state-level achievement in 2022-23 may have been 0.010 SDs higher if absence rates had returned to pre-pandemic norms in this year. Increasing test scores by this amount would make up 6% of the 0.176 SD gap in test scores that remained between students in 2022-23 and their peers in 2018-19.

Second, we find that the relationship between absenteeism and achievement is weaker post-COVID than pre-COVID. In our pre-COVID cohorts, each absence is associated with a 0.0055 SD reduction in achievement, consistent with estimates from other pre-COVID studies in this setting (Aucejo & Romano, 2016; Gershenson et al., 2017). These estimates imply that attendance recovery in 2022-23 might have improved test scores in this year by 0.018 SDs.

We additionally estimated models that include peer absence effects. We find null effects of peer absenteeism post-pandemic, while results are mixed pre-pandemic, with the relationship ranging from null to a 0.0035 SD decrease per one-day increase in the average absence rate of classroom peers. Depending on assumptions about peer effects, pre-COVID models suggest that returning attendance to pre-COVID levels in 2022-23 would have improved achievement in this year by between 0.018 and 0.031 SDs, or 10% to 18% of the total needed to achieve a full recovery. In summary, across all models (pre- and post-COVID and with and without peer effects) our estimates range from suggesting that returning attendance to pre-pandemic levels in 2022-23 could have improved math test scores by 0.010 to 0.031 SDs, or 6% to 18% of the total needed to achieve a full recovery.

These results point to three key takeaways. First, as anticipated based on pre-pandemic research, increased post-pandemic absenteeism is likely slowing academic recovery. As such, investments to improve attendance rates will likely have positive spillover effects on achievement. However, second, recovering attendance rates would not fully recover achievement. Therefore, more direct efforts to remediate learning loss will still be necessary. Finally, the relationship between absenteeism and achievement has become weaker post-pandemic. Future research should aim to confirm this result in other settings and identify factors underlying this change.

II. Literature Review

a) *Impacts of COVID-19 on educational outcomes*

In March 2020, schools across the U.S. shut down and switched to providing only virtual instruction due to the COVID-19 pandemic. In 2020-21, most schools continued to operate virtually in the beginning of the year, but increasingly moved to hybrid or fully in-person instruction as the year progressed, including in this study's setting of North Carolina (COVID-19 School Data Hub, 2022). By 2021-22, most public schools, including those in North Carolina, had returned to full in-person instruction (Granados, 2021).

The disruptions caused by the pandemic had dramatic effects on educational outcomes. Student achievement declined substantially between 2018-19 and 2020-21 (e.g., Goldhaber et al., 2023; Kuhfeld et al., 2022). Since 2020-21, there has been only partial academic recovery, with test scores remaining well below pre-pandemic levels as of 2022-23, especially among Black and Hispanic students, students in poverty, and students with the lowest baseline achievement (Fahle et al., 2022, 2024; Lewis & Kuhfeld, 2023; NCES, 2022, 2024; Peters et al., 2023). In addition, some research suggests that recovery stagnated in 2022-23 and 2023-24, with students demonstrating only similar or less growth than in a typical pre-pandemic year (Curriculum Associates, 2023; Kuhfeld & Lewis, 2024; Lewis & Kuhfeld, 2023) – though other research finds that there has been accelerated growth on state exams (Fahle et al., 2024). Overall, Lewis and Kuhfeld (2023) estimate that students in 2022-23 scored about 0.16 to 0.27 SDs below their pre-pandemic peers on MAP math assessments, equal to 4.5 months of schooling on average.

Increases in student absenteeism have also emerged as a key post-pandemic problem. Nationally, the percentage of students who were chronically absent – defined as missing 10% or more of school days, or about 18 days in a typical school year – increased from 15% in 2018-19 to 28% in 2021-22 (Dee, 2024; Malkus, 2024). Since then, absence rates have declined only a little (Malkus, 2024). North Carolina, the setting of this study, has experienced similar changes (Fuller et al., 2024). Further, in North Carolina, research shows that 40% of elementary and middle school students were chronically absent at least once between 2020-21 and 2022-23, while 8% were chronically absent all three years (Swiderski et al., 2024). These rates are 2.5 to 3 times higher than in a comparable pre-pandemic period.

While chronic absenteeism has increased, it is also important to note that the entire absence distribution has shifted upward. For example, in North Carolina, the median student missed 7 days of school in 2018-19, but 11 days in 2021-22 and 10 days in 2022-23 (Fuller et al., 2024). This shows that absence rates have not just increased among highly-absent students, but that many students are missing at least a few more days of school than usual each year.

b) The relationship between absenteeism and achievement

Prior research consistently finds a negative relationship between student absenteeism and achievement (Ansari & Gottfried, 2021; Aucejo & Romano, 2016; Gershenson et al., 2017; Goodman, 2014; Gottfried, 2011; Gottfried & Kirksey, 2017; Liu et al., 2021; Wei, 2024). The studies most closely related to ours in terms of setting, population, and method are those of Gershenson et al. (2017) and Aucejo and Romano (2016), each of which studied the impact of absenteeism on achievement among late elementary school students in North Carolina using student fixed effects or value-added models. Gershenson et al. (2017) estimate that each additional absence decreased student achievement in math by 0.007 SDs when employing a value-added model and 0.005 SDs when employing a first-differences model. Aucejo and Romano (2016) also estimated about a 0.006 SD decline in math achievement per additional absence in their preferred specification. These results suggest that increased post-pandemic absenteeism may be slowing the pace of academic recovery.

In addition, some research also suggests that the absence rate of classroom peers may have negative spillover impacts on a student's own achievement (Gottfried, 2019; Gottfried & Ansari, 2022; Monk & Ibrahim, 1984). This might be because widespread absenteeism can disrupt instruction. For example, when many students miss many days of school, teachers may need to spend more time reviewing material or working with the students who have been absent to help them catch up. This suggests that the population-wide increases in absenteeism post-pandemic may be damaging to learning recovery beyond the individual effects of each student's own increase in absenteeism.

c) *The current study*

Although there is consistent evidence that absenteeism is negatively associated with achievement, we are aware of only one study that has examined how the increase in post-pandemic absenteeism may be related to achievement declines, which found that increased self-reported absenteeism in the month leading up to the NAEP exam could explain 16% to 27% of changes in NAEP math scores (NCES, 2023). In the current study, we expand on this work by estimating the relationship between absenteeism and achievement in a large state setting using detailed individual-level administrative data.

We undertake this study in part because while prior literature consistently identifies a negative relationship between absenteeism and achievement, magnitudes vary across studies and settings. Moreover, it is possible that the relationship between absences and achievement has changed over time. For example, some reports suggest that school materials are more widely available online than pre-pandemic (Jacob & Stanojevich, 2024; Mervosh & Paris, 2024; The Learning Network, 2024), which might help students keep up with schoolwork from home more easily and thus make each day missed less costly than pre-pandemic. Relatedly, absenteeism might also be less costly if instructional pacing has slowed, potentially due to a continued need to remediate students from pandemic learning losses or from disruption created by the classroom-wide increases in absenteeism. On the other hand, absences might be more costly than pre-pandemic as schools are working to accelerate student learning. For example, being absent might mean missing out on school-day or after-school tutoring programs that have been widely put in place as a recovery effort (National Student Support Accelerator, 2023). For these reasons, it is unclear as to whether the post-pandemic impact of absenteeism on achievement will be the same as the pre-pandemic impact.

III. Data

We use individual-level administrative data from the North Carolina Department of Public Instruction spanning the 2015-16 through 2022-23 school years, including student demographics, attendance and suspension records, classroom rosters, and state test scores. We focus on students in grades 3 through 7, which are the grade levels where all students take end-of-grade exams in math.[1]

Our outcome is a student's standardized end-of-grade math test score, which we derive from students' exam scale scores. To allow us to measure pandemic-induced changes in achievement, we anchored post-COVID scale scores to the 2018-19 distribution. For example, a student who received a scale score of 548 in 6th grade math in 2018-19 had a standardized score of -0.016 in this year; therefore, we assigned a score of 548 in 6th grade math to have a standardized score of -0.016 in post-COVID years as well.[2]

Our key treatment variable is the number of days a student was absent. Because students are not all observed for the same number of days each year, we normalized absence totals to a 180-day year by multiplying a student's percent of days absent by 180. For example, a student who was absent for 5 days out of 120 days enrolled (4.2% of days) would have a normalized absence total of 7.5 days.[3]

[1] Students who enroll in high school Math 1 in 8th grade take the Math 1 end-of-course exam in place of their 8th grade end-of-grade exam in North Carolina. This also applies to students in lower grades who enroll in Math 1, but enrollment in this course before 8th grade is rare (e.g., less than 3% of 7th graders). To be more specific, to be included in our sample, students needed to have taken a grade 3 through grade 7 end-of-grade math exam.
[2] Scale scores are integers that range from about 520 to 580 regardless of grade level in 2018-19 and 2022-23 and from about 420 to 480 prior to 2018-19. Scale score ranges changed in 2018-19 due to a new version of the test being implemented. We standardized pre-pandemic scores within test year to have a mean of 0 and standard deviation of 1. An increase of 1 scale-score point is equivalent to about a 0.10 SD increase. Because of the change in the exam scale, we cannot anchor scores from prior to 2018-19 to the 2018-19 distribution. We tested sensitivity to using within-year rather than anchored standardized scores as the outcome for post-pandemic cohorts and found no substantive difference in the results, as described more in our Sensitivity Analyses section below.
[3] This adjustment affects most students by a small amount, as the attendance data we receive are from shortly before the end of the school year, such that most students have 165-175 days in membership at the time of data collection.

Figures 1 and 2 depict trends in the distribution of these variables between 2018-19 and 2022-23. Figure 1 shows the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of math test scores in each year. In general, all math achievement percentiles dropped substantially in 2020-21, recovered partially in 2021-22, and recovered further (yet modestly) in 2022-23. With the exception of the 10^{th} percentile, the initial impact of the pandemic was more negative at lower percentiles, and recovery in 2021-22 and 2022-23 was generally less at the bottom of the distribution than at the top. As a result, as of 2022-23, the 90^{th} percentile was only 0.03 SDs below its 2018-19 value, whereas the 75^{th} percentile remained 0.12 SDs below, the median remained 0.14 SDs below, and the 25^{th} percentile remained 0.29 SDs below. The exception to this trend is that the pandemic's initial impact at the 10th percentile was smaller than at other percentiles, which contrasts results from national data (NCES, 2022; Peters et al., 2023). We suspect that this may be due to floor effects on the state exams constraining the drop at the bottom of the distribution.

Figure 2 shows the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles of days absent. In 2020-21, when schools operated with a mix of remote, hybrid, and fully in-person instruction, days absent dropped at the 10^{th}, 25^{th}, and 50^{th} percentiles, increased modestly at the 75^{th} percentile, and increased substantially at the 90^{th} percentile. However, attendance data in this year might be inconsistent with other years or across contexts due to possible differences in the measurement and tracking of attendance across instructional modalities. Further, what it means to have attended school, largely virtually, during this year may have been different than other years. In 2021-22, when schools had returned to fully in-person learning, the absence distribution was strictly higher than in 2018-19, with the greatest increases at the top of the distribution (though the 90^{th} percentile had decreased relative to 2020-21). Between 2018-19 and 2021-22, the 10^{th} percentile increased from 1 to 2 days absent, the median increased from 7 to 11 days, and the 90^{th}

percentile increased from 19 to 30 days. Finally, between 2021-22 and 2022-23, absenteeism declined at the top and middle of the distribution (e.g., from 30 to 27 days at the 90th percentile and 11 to 10 days at the median) but was steady at the bottom of the distribution.

IV. Method

a) Primary models

Our primary aim is to estimate the relationship between absenteeism and achievement. A standard regression of current-year achievement on current-year absences and other observable controls would likely be biased by unobservable factors that correlate with both a student's attendance and their achievement. We therefore employ first-differences models, which are equivalent to student fixed effects models for the two-period case, to net out the impact of time-invariant unobservable factors. These models take the form:

$$(1)\ Y_{it} - Y_{it-1} = \beta_0 + \beta_1(Abs_{it} - Abs_{it-1}) + \beta_2(X_{it} - X_{it-1}) + \epsilon_{it},$$

where Y_{it} refers to student i's math test score in year t, Abs_{it} refers to their days absent, X_{it} is a set of time-varying student-level characteristics, and ϵ_{it} is an error term. The key term is β_1, which provides an estimate of the effect of absenteeism on achievement. We additionally use β_1 to generate a back-of-the-envelope estimate of what test scores in 2022-23 might have been had absence rates returned to pre-pandemic levels in this year, which we obtain by multiplying β_1 by the average increase in days absent from pre- to post-pandemic. In our main specification, we limit X_{it} to a single variable – the number of days a student was suspended, which is a covariate that might meaningfully vary from year to year and correlate with both absenteeism and achievement.[4] We do not necessarily assume that this variable is unconfounded and therefore do not interpret the β_2 coefficient in the results below, though we do present it for reference.

We estimate equation (1) for t = 2016-17, t = 2017-18, t = 2018-19, and t = 2022-23, thus providing three pre-pandemic estimates and one post-pandemic estimate. We compare β_1 terms across the pre-pandemic and post-pandemic cohorts to identify whether the relationship between absenteeism and achievement has changed (based on whether standard errors of β_1 estimates overlap). We omit 2019-20 and 2020-21 because no exams were taken in 2019-20. We omit 2021-22 from our main analyses out of concern that attendance data in 2020-21, when schools operated with a mix of remote, hybrid, and fully in-person instruction, may not be comparable across modalities or to other years, which could lead to measurement error that would bias results. We address this concern more in the Sensitivity Analyses section.

Equation (1) provides a causal estimate of the effect of absenteeism on achievement under the assumption that there are no time-varying confounders that are not accounted for in our model. While this is a more robust assumption than the assumptions of a cross-sectional regression model (i.e., that there are no time-varying or time-invariant confounders), it could still be violated. For example, it may be that some students or their parents become more (or less) engaged with school between years, leading a student to both attend school more and improve their test scores. This untestable assumption is a key limitation of this study.

We apply several sampling restrictions to our primary analyses. First, we omit students who took math in a virtual school, as the relationship between absenteeism and achievement may

[4] Other observable data points, such as English Learner status, can vary over time within an individual. However, we omit other characteristics under the assumption that changes may be somewhat arbitrary due to the nature of how such variables are measured. For example, some students may switch from 1 to 0 on this variable by moving from being almost proficient to fully proficient, whereas students who move from knowing no English to being somewhat proficient might show no change in this status. Ultimately, we tested sensitivity to including a larger set of controls and found that this was inconsequential to the estimates, as described in the Sensitivity Analyses section.

differ in these environments.[5] We also omit students in classrooms identified as having 4 or fewer or 40 or more total students, which might also represent virtual or other atypical learning environments, as well as those who did not appear in classroom roster data (identification of classrooms is described more below). Combined, these restrictions exclude just under 4% of student-year observations between 2015-16 and 2022-23. We then omit 3% of remaining students who had fewer than 90 days enrolled, more than 50% of days absent or suspended, who were grade retained, or whose tested grade level did not match their school-reported grade level to reduce skew due to outliers, errors, or other unique circumstances. An additional 3% of remaining students were missing test score, absence, or suspension data and were excluded. Finally, we restrict our main analyses to students whose baseline achievement was in the top 80% of the achievement distribution to reduce bias due to possible floor effects on the state exams, as informed by Figure 1.[6]

Table 1 provides descriptive statistics of the analytic sample. Variables at the top of the table provide statistics about the composition of students in each year. Due to our restriction of the sample to students in the top 80% of the achievement distribution, the average standardized math score is above 0, with mean standardized scores of 0.31 to 0.32 pre-pandemic and 0.14 in 2022-23. Average days absent pre-pandemic ranged from 7 to 8, but reached 11.3 in 2022-23. Students in 2022-23 also averaged more days suspended than pre-pandemic peers, with a mean of 0.66 in 2022-23 as compared to 0.43-0.48 pre-pandemic. Across all years, about half of

[5] This is defined as students in a school with "Virtual" in the title as well as students in the NC Cyber Academy.
[6] "Baseline" refers to the first year of the two-year period used to estimate the model.

students are White, 11% are English Learners (except in 2016-17), 6-7% have a disability, 19% are academically or intellectually gifted, and just under half are in urban schools.[7]

The bottom of Table 1 shows the three first-differenced variables used in Equation (1). Overall, students in our sample tended to experience a slight decrease in math scores from the prior year, which likely represents regression to the mean arising from our sample restriction to students in the top 80% of the baseline achievement distribution. Changes in days absent were slightly positive in pre-pandemic cohorts, but slightly negative in 2022-23. Finally, changes in days suspended are positive in all cohorts (as students moved into higher grade levels), with the increase being slightly above pre-pandemic norms in 2022-23.

Appendix Table A1 provides more detail about the distribution of the first-differenced treatment and outcome variables (see also Appendix Figure A1 for a histogram of these distributions). There was more dispersion in the first-differenced absence variable post-pandemic relative to pre-pandemic. For example, in 2018-19, the 5th percentile change was a decrease of 8 absences from the prior year; in 2022-23, it was a decrease of 14 days. Meanwhile, the 95th percentile change in 2018-19 was an increase of 10 days, whereas the 95th percentile in 2022-23 was an increase of 12 days. In terms of achievement, the 5th to 95th percentile change in 2018-19 ranged from a decrease of 0.94 SDs to an increase of 0.80 SDs. In 2022-23, the range was from a decrease of 0.90 SDs to an increase of 0.90 SDs.

[7] We do not include students' economic disadvantage status because this variable was significantly affected by COVID-era universal free lunch policies, which reduced the percentage of students who were recorded as being economically disadvantaged, likely because families did not have to complete free/reduced-price lunch forms to qualify for free lunches. For example, in 2021-22, only 40% of students were classified as economically disadvantaged, compared to 49% in 2018-19 and 53% in 2022-23.

b) Peer effects models

In addition to the primary models above, we also tested models that include peer absence rates as a covariate. We identified a student's math classroom peers using classroom roster data. We considered students to be in the same classroom if they were enrolled in the same *section* of the same core *math course* (based on a 10-digit alpha-numeric code) taken during the same *term* in the same *school*. Consistent with the main sample, we omitted students from classrooms if they had fewer than 90 days enrolled, more than 50% of days absent or suspended, or were enrolled in a virtual school.

We identify the absence rate of classroom peers as the (normalized) average days absent of students in the classroom excluding the focal student.[8] In these models, we also control for the percentage of peers who were suspended and the average prior math test score of classroom peers. Because of the latter covariate, we can estimate these models only for students in grades 4 through 7 (i.e., because grade 3 is the first tested grade, the classroom peers of students in grade 3 do not have a prior test score average) and for t = 2017-18, 2018-19, and 2022-23. Appendix Tables A2 and A3 and Appendix Figure A2 show descriptive statistics pertaining to the classroom peer variables included in these models. In 2022-23, students were in classrooms with peers who averaged more absences (12.4) than pre-pandemic (8 to 9) and where a higher proportion of peers were suspended (17% vs. 14%). Peer absence rates are higher than individual absence averages because students in the bottom achievement quintile are included when calculating classroom averages, but are excluded from the analytic sample, and these students tend to have higher absence rates. In terms of first-differences, students in 2022-23 tended to

[8] About 3% of students appear in more than one math class in a single year. For these students, we use the average of their classroom averages.

experience a decrease in peer absenteeism relative to the prior year, whereas pre-pandemic peers tended to experience an increase in peer absenteeism across years.

V. Results

a) Mean changes in absenteeism and achievement

Table 2 presents results from cross-sectional regression models that estimate changes in mean standardized math test scores, days absent, and peer days absent in 2022-23 relative to 2018-19, which guide our back-of-the-envelope calculations below. Panel A shows models that are restricted to the analytic sample, while Panel B shows models that include students in the bottom baseline achievement quintile. Models 1, 3, and 5 include no covariates and are thus equivalent to descriptive changes. Models 2, 4, and 6 adjust for the demographic variables shown in Table 1.[9] In all cases, adjusted and unadjusted results are nearly identical.

In the analytic sample, results show that students' test scores in 2022-23 remained about 0.176 SDs lower than their peers from 2018-19. Meanwhile, students in 2022-23 averaged about 3.27 additional days absent. Finally, students in 2022-23 were in classrooms where peers averaged about 3.65 additional days absent. For reference, we note that the decline in test scores is slightly less (about 0.163 SDs) when students in the bottom quintile are included, while increases in absences (about 3.82 days) and peer absences (3.90 days) are slightly greater.

b) Correlations between absenteeism and achievement

As a precursor to the main analytic results, Figure 3 plots the simple bivariate correlation between absenteeism and achievement in 2016-17, 2017-18, 2018-19, and 2022-23 (restricted to students in the analytic sample). This plot highlights three key points. First, as expected, absenteeism is negatively correlated with achievement in each year. Second, students in pre-

---

[9] That is, we estimate: $Y_{it} = \beta_0 + \beta_1 Year_{it}(+\beta_2 X_{it}) + \epsilon_{it}$, where $Year_{it}$ = {2018-19, 2022-23}.

pandemic years tended to score higher than peers in 2022-23 who had the same number of absences. However, third, the magnitude of the negative correlation between absenteeism and achievement is approximately the same in all years (see also Appendix Table A4).

c) *First-differences estimates of the relationship between absenteeism and achievement*

Table 3 presents the core results of this study. Models 1 through 4 show results from first-differences models with only the days absent variable (note that tables show estimates pertaining to a 10-day increase in absenteeism rather than a 1-day increase because of the small magnitude of the coefficients and standard errors). Across the three pre-pandemic cohorts (Models 1 through 3), the coefficient is extremely stable. In each case, each day absent is associated with a 0.0063 SD decrease in achievement. These results are similar to prior estimates from North Carolina (Aucejo & Romano, 2016; Gershenson et al., 2017). However, post-pandemic (Model 4), we identify a weaker, though still significantly negative, relationship, with each day absent associated with a 0.0040 SD decrease in achievement.

Models 5 through 8 repeat this analysis including days suspended as a covariate (these results are also depicted in a coefficient plot in Figure 4, while Appendix Figure A3 presents a plot of predicted scores in each year as a function of days absent). Across models, including this covariate reduces the estimated effect of an additional absence, but the pattern of change over time remains qualitatively the same. Pre-pandemic, each day absent is associated with about a 0.0055 SD decrease in achievement in these models. However, in 2022-23, each day absent is associated with a 0.0032 SD decrease in achievement. The coefficient on days suspended is similar across all four models.

Overall, students in our analytic sample averaged 3.27 additional days absent in 2022-23 compared to 2018-19. Using the post-pandemic estimate from Model 8, a back-of-the-envelope estimate suggests that returning attendance to pre-pandemic levels in 2022-23 may have resulted in achievement being 0.010 SDs higher in this year (3.27 days x 0.0032 SDs per day). This would make up 6% of the 0.176 SDs needed to achieve a full academic recovery. Meanwhile, the pre-pandemic estimate from Model 7 suggests that returning to pre-pandemic attendance rates may have increased achievement by 0.018 SDs in this year, or 10% of the total needed for full recovery.

d) Peer effects models

Table 4 presents results from models that include peer effects. Models 1 through 3 show estimates from models that only include own and peer absences, while Models 4 through 6 include controls. All models produce similar estimates of the effect of one's own absenteeism as corresponding models from Table 3. Within the pre-pandemic period, we do not consistently identify a significant negative association between peer absenteeism and a student's own achievement. In Model 4 (the 2017-18 cohort), we estimate that the impact of peer absenteeism is directionally negative but null, whereas in Model 5 (the 2018-19 cohort), we estimate that a one-day increase in the average absence rate of classroom peers reduces a student's own achievement by a statistically significant 0.0035 SDs. Post-pandemic, Model 6 shows that we identify a null, near-zero relationship between peer absenteeism and achievement.

Given that peer effects may be null, and that adding peer effects did not substantively change estimates of the impact of one's own absenteeism on achievement, these models do not necessarily affect our back-of-the-envelope calculations of the potential impact of recovering attendance. However, using the estimates from Model 5 where significant negative peer effects in the pre-pandemic period were identified, we would estimate that full attendance recovery may have improved achievement in 2022-23 by 0.031 SDs, or 18% of the total needed to achieve a full recovery.[10]

e) Sensitivity analyses

We next test the sensitivity of these results to several alternative ways of defining the sample, key variables, or model. We display results for the 2018-19 and 2022-23 cohorts, with Appendix Table A5 showing results for models with individual effects only and Appendix Table A6 displaying results for models that include peer effects.

Focusing on Table A5, Model 1 replicates the main estimates for reference. In Model 2, we test sensitivity to including additional time-varying covariates and a school fixed effect. These results confirm that including additional covariates does not affect our estimates.

In Model 3, we return to our main specification (no additional covariates beyond those included in the main model) but include a squared days absent term. This term is statistically significant and negative pre-pandemic, suggesting that each additional day absent has a more negative relationship with achievement at higher levels of absenteeism, but is null post-pandemic. In either case, the magnitude is small, consistent with prior research that suggests that the relationship between absenteeism and achievement is approximately linear (Gershenson et al., 2017). Adding this term does not change the conclusion that the relationship between absenteeism and achievement has weakened over time.

Model 4 tests sensitivity to using within-grade, within-year standardized math scores as the outcome for the post-COVID cohort rather than the anchored standardized score. Model 5 tests sensitivity to including additional sample restrictions where we omit a small percentage of students who experienced outlier changes in either their absence rate or achievement from the

[10] This is calculated as 3.27 days x 0.0055 SDs per own day absent plus 3.65 days x 0.0035 SDs per classroom peer average days absent.

prior year (a change of more than 25 absences or 1.3 SDs in achievement), which checks whether main results might be distorted due to outliers.[11] Model 6 tests sensitivity to excluding charter schools. All of these estimates are very similar to the main estimate.

In Model 7, we test sensitivity to including students in the bottom baseline achievement quintile. Including these students weakens the pre-pandemic estimate by about 20% to a 0.0044 SD reduction in achievement per day absent. Including these students has little impact on the post-pandemic estimate, which remains at a 0.0033 SD reduction per day absent.

Finally, in Model 8, we test sensitivity to using a value-added model (VAM). To conduct the VAM, we regressed a student's test score in year t on their test score in year $t - 1$ and other observable covariates (including days absent) in year t.[12] Similar to other studies that have employed both VAM and first-differences approaches (Aucejo & Romano, 2016; Gershenson et al., 2017), we find that, as compared to the first-differences model, the VAM produces slightly more negative estimates of the effect of absences on test scores. However, VAMs continue to show a weakened relationship between absenteeism and test scores in 2022-23 relative to 2018-19, with the magnitude of the change being only slightly less than the magnitude of change in the first-differences specification.[13] We privilege the first-differences estimates in this manuscript as they produce more conservative estimates of the effect of absences on test scores.

---

[11] Within the peer effects model in Appendix Table A6, we also omit students experiencing an outlier change in peer absences (a change of more than 15 days in average peer absences).

[12] This approach relies on relies on variation in absences *between* students in year *t* rather than *within* students across time for identification. For a full discussion of differences between VAM and first-differences models, see Angrist & Pischke (2009). The complete set of observed covariates used in the own-absence model include: student grade-level, race/ethnicity, sex, economic disadvantage status, English Learner status, disability status, days suspended, and an indicator of having attended multiple schools; and a school fixed effect. The peer-absence VAM model additionally includes: math classroom peer-level prior math scores, suspensions, sex, race/ethnicity, economic disadvantage status, English Learner status, disability status; and the number of students in the math class.

[13] Like the first-differences models, VAM specifications also identify a similar effect of absenteeism in 2016-17 and 2017-18 as 2018-19 and that the effect of absenteeism decreased from pre- to post-pandemic within all subgroups examined. At times, some patterns of relationships across subgroups differ from the main model results (described below) – e.g., VAM specifications suggest that absenteeism may be somewhat more harmful for high-achieving than

Table A6 presents these sensitivity checks for models that include peer effects. Findings are largely the same, except that pre-pandemic peer effects become null in Model 7. In addition, VAMs produce slightly weaker estimates of peer effects than the main model in 2018-19, but more (and statistically significantly) negative in 2022-23.[14] As a result, VAMs do not affect the bounds of our back-of-the-envelope estimates of how test scores might have changed in 2022-23 had attendance rates returned to pre-pandemic levels. Using the VAM from 2018-19, we would estimate that returning attendance to pre-pandemic norms would have improved test scores by 0.022 SDs in 2022-23 if we do not consider peer effects (Table A5) and by 0.030 SDs if we do (Table A6). Using the VAM from 2022-23, we would estimate that test scores would have increased by 0.016 SDs if we do not consider peer effects and by 0.029 SDs if we do. These estimates all fall within the bounds of the back-of-the-envelope estimates produced by the first-differences models, which ranged from 0.010 to 0.031 SDs.

Finally, Appendix Table A7 shows results for the 2021-22 cohort. This cohort is excluded from our primary estimates due to concerns that attendance data in 2020-21 may not be comparable to other years, or consistent across contexts, due to the mix of remote, hybrid, and in-person instruction in this year, which could produce measurement error that could attenuate first-differences results towards zero. Indeed, Table A7 shows that we identify almost no relationship between absenteeism and achievement in this cohort using our first-differences model. However, we also produced results for 2021-22 using a VAM. As the VAM does not rely on absence data from 2020-21, it surmounts concerns about measurement error in this year. The VAM produces a negative effect estimate in 2021-22 that is similar to the effect estimate it

---

middle-achieving students. However, as key conclusions are similar, for simplicity, we show only VAM specifications related to the main analysis in this manuscript. Full VAM results are available on request.

[14] A small number of students are missing covariates that are used in the VAM but not in the first-difference model, resulting in a marginally reduced sample size.

produces for 2022-23. We suggest that this: 1) shows that concerns about measurement error of attendance in 2020-21 may be valid; 2) suggests that there was likely a negative effect of absenteeism in 2021-22; and 3) bolsters the conclusion that there has been a shift in the relationship between absenteeism and achievement from pre- to post-COVID, as VAM estimates for 2021-22 and 2022-23 are both similar to each other and weaker than pre-pandemic.

f) Subgroup analyses

Finally, we present results from subgroup analyses disaggregated by student demographics and baseline achievement. In each case, we obtain estimates by re-running Equation (1) restricted to members of the focal subgroup. For simplicity, we focus on the 2018-19 and 2022-23 cohorts only. Further, because of the inconsistency in peer effect estimates in the pre-pandemic period, and because excluding this variable did not affect estimates of a student's own absenteeism on their achievement, we focus only on models that do not include peer effects.

Table 5 presents results by race/ethnicity. Appendix Table A8 additionally shows results by grade level, gender, and disability status, while Appendix Table A9 shows results that include peer effects for race/ethnicity and grade level subgroups. In general, differences between student subgroups are modest within each cohort. For example, results by race/ethnicity range from 0.0050 to 0.0057 SD impacts pre-pandemic and 0.0032 to 0.0039 SD impacts post-pandemic. Further, within all subgroups, the relationship between absenteeism and achievement weakens across time from pre- to post-pandemic. These results suggest that there is little heterogeneity in the relationship between absenteeism and achievement and that changes over time in the main sample are not due to changes in characteristics of the student population, as all student subgroups experienced similar changes.

Table 6 presents results disaggregated by students' baseline achievement quintile, including students in the bottom quintile who were omitted from the main analytic sample. For comparability across time, we place post-COVID students into academic quintiles based on the 2018-19 distribution. This ensures that, for example, students in quintile 1 in 2022-23 have similar scale scores as those in quintile 1 in 2018-19. Doing so also means that, in 2022-23, more students are in the bottom "quintile" than the top. An alternative where students in 2022-23 are placed into quintiles based on their position within their own baseline (2021-22) distribution is presented in Appendix Table A10, which produces similar results.

Results in Table 6 show that, in 2018-19, the negative relationship between absenteeism and achievement was strongest among middle-achieving students (a decrease of 0.0063 SDs per additional day absent), whereas we identify weaker associations for students at the bottom (0.0023) and top (0.0048) of the distribution. However, across all achievement subgroups, estimated impacts are weaker in 2022-23 than 2018-19, suggesting again that the change over time in the full sample is not driven by changes in students' baseline proficiency. In 2022-23, the pattern of estimates across achievement subgroups is generally the same as in 2018-19, except that we find an equally large association for students in the top quintile as those in the middle.

Overall, subgroup analyses do not reveal substantial differences in the relationship between absenteeism and achievement based on demographic characteristics, though there may be some variation across achievement subgroups. However, within all subgroups, we identify a substantial weakening of this relationship between 2018-19 and 2022-23.

VI. Discussion

With student achievement remaining significantly below pre-pandemic levels, there are many questions about how to enhance academic recovery. Increased student absenteeism has emerged as a key point of concern among researchers and policymakers, and prior research suggests that this could be slowing academic recovery.

The estimates in this study suggest that, pre-pandemic, each additional day absent was associated with a 0.0055 SD decrease in achievement, similar to prior estimates in this context (Aucejo & Romano, 2016; Gershenson et al., 2017). We also identified mixed evidence of a negative relationship between the absence rate of classroom peers and a student's own achievement. Back-of-the-envelope estimates from pre-pandemic models suggest that returning attendance rates to pre-pandemic norms in 2022-23 would have resulted in achievement being about 0.018 SDs to 0.031 SDs higher in this year, which would account for 10% to 18% of the total 0.176 SDs needed to achieve a full achievement recovery among the students in our analytic sample. This range overlaps with one prior study that found that increased absenteeism may have accounted for 16% to 27% of the decreases in NAEP test scores (NCES, 2023).

However, we identify a substantially weaker relationship between absenteeism and achievement post-pandemic. In 2022-23, we find that each additional day absent was associated with a 0.0032 SD decline in achievement and find null effects of peer absenteeism. This result suggests that attendance recovery in 2022-23 might have only resulted in achievement being 0.010 SDs higher than observed, or 6% of the total needed for full achievement recovery.

This range of estimates – from 6% to 18% – is relatively wide, reflecting a considerable decline in the estimated impact of a student's own absenteeism and uncertainty about the impact of peer absenteeism. However, it is not clear that post-COVID estimates are strictly preferable when it comes to our back-of-the-envelope calculations. For example, if the causes underlying the change in the estimated impact of absenteeism are temporary and/or caused in part by the increase in absenteeism, then we might prefer pre-pandemic estimates.

It is therefore important to consider why the impact of absenteeism on achievement may have changed. To begin, we remind readers that there are limitations of our data and method. We observe only one post-pandemic cohort of students in one state and rely on the assumption that there are no time-varying variables that affect both absenteeism and achievement that we do not observe in our model. It is therefore possible that an idiosyncrasy of our data or a (more or less severe) violation of our methodological assumption is the true cause of these changes over time. Relatedly, the change could be due to changes in unobservable characteristics that we cannot control for. For example, it could be that the timing of absenteeism has changed in ways that affect these results. Absences accrued near a test date have more impact on achievement than more distal absences (Gottfried & Kirksey, 2017; Liu et al., 2021). Increases in post-pandemic absenteeism that are driven by increased seasonal illness in the Fall or Winter – or a willingness to miss school due to seasonal illness – might have a limited impact on student end-of-grade test scores and thus attenuate estimates of the impact of total yearly absences.

However, there have also been changes in the context and nature of schooling that might explain these results. For one, some reports suggest that teachers are more often posting course material online, which may make it easier for students to keep up with work from home and thus make absenteeism less costly (Jacob & Stanojevich, 2024; Mervosh & Paris, 2024). While this might induce some students or families to become more lax about school attendance (The Learning Network, 2024), this would also have a positive side – students will always miss at least some school due to sickness, family emergencies, and other events, so making these absences less costly would be beneficial, on average, if students do not offset these benefits by missing more school. In other words, this mechanism might reduce the costliness of absenteeism by increasing the rate of learning during or following an absence. This type of change might also be permanent, so long as teachers continue to make resources available online.

Other potential explanations are more transient. For example, absenteeism may have become less costly because the pace of instruction is currently slower as teachers continue to remediate students from pandemic learning losses and deal with widespread absenteeism. If so, then it is possible that as achievement and attendance rates return to pre-pandemic norms, instructional pacing might also increase, and absences would once again become as costly as pre-pandemic. In the short-term, however, this mechanism would mean that absenteeism may have been temporarily less costly due to a lower overall rate of learning.

An additional possibility is that the increased availability of tutoring and other efforts to provide students with individualized instruction may have mitigated the costliness of absenteeism by providing students with more resources to catch back up once they returned to school. Like our first mechanism, this implies that the costliness of absenteeism may have changed due to changes in students' rate of learning during or following an absence, with no implication that there was a change in students' rate of learning when present in school. Like our second possibility, this mechanism might also be temporary, as schools may begin to cut tutoring services as COVID relief funding runs out and as student achievement recovers.

While we cannot distinguish between these possibilities with the current data, this discussion highlights that it is not clear that the negative impact of absenteeism has permanently weakened, nor should results be taken to mean that school *attendance* has become less important to achievement. However, taken together, our results raise the possibility that absenteeism may be higher in part because the costliness of absenteeism is lower than pre-pandemic. Consider, for example, that our 2018-19 model implies that a student who missed an average number of days of school in that year (8.04) would have achieved a test score 0.044 SDs lower than if they had had perfect attendance in that year. Meanwhile, in 2022-23, a student who missed an average number of days in that year (11.31) would have achieved a test score just 0.036 SDs lower than if they had had perfect attendance in that year.[15] Thus, our results imply that students may still be relatively better-off academically in 2022-23 than they were in 2018-19 even with an increased absence rate.

We note that we consider only the relationship between absenteeism and achievement and only in one subject. Absenteeism may have other consequences, such as weakening a student's relationships with peers or teachers (Gottfried et al., 2024). Additionally, the results of this study focus on the 3.3-day increase in absenteeism that occurred on average. However, a return to a pre-pandemic absence distribution would affect some students more than others. At the low end of the distribution, students may be missing only 1 to 2 more days of school than usual; at the high end, students are missing several weeks more school than usual. Results from Table 2 suggest that students with the lowest baseline achievement are also experiencing the highest increases in absenteeism, but we are uncertain about whether we can observe the full extent to which these students' achievement has dropped due to possible floor effects on these state exams, meaning that our estimates do not take these students – and the possible gains from recovering these students' attendance – into account. Improving attendance rates among the highest-absence students might yield higher returns while also helping to narrow the achievement distribution back to pre-pandemic norms.

Finally, it is important to note that we take the increased absence rate in 2021-22 as a given, with our back-of-the-envelope estimates pertaining only to what might have happened if

[15] The estimate for 2018-19 is calculated as (8.04 days x -0.0055 SDs per absence). The estimate for 2022-23 is calculated as (11.31 days x -0.0032 SDs per absence).

attendance had recovered in 2022-23. However, research shows that the effects of absenteeism may accumulate across time (Ansari & Gottfried, 2021; Wei, 2024). If absenteeism had returned to pre-pandemic norms in 2021-22 and remained that way in 2022-23, test scores in 2022-23 might be substantively higher than our back-of-the-envelope estimates suggest.[16] A sustained increase in absences may, over time, come to account for increasing portions of the achievement gap between pre- and post-COVID students.

Overall, this study finds that student absenteeism may have been less costly in 2022-23 than pre-pandemic, but was still associated with statistically significant decreases in student achievement. These results suggest that attendance recovery can play a key role in accelerating learning recovery. However, achieving a full academic recovery will require more than just attendance recovery. Given the relatively large change in the magnitude of our estimates across cohorts as well as the limitations of our data and method, we encourage replication of this research in other settings, particularly as additional post-COVID cohorts can be observed. We also encourage exploration of how the ways in which students, teachers, and schools respond to absenteeism have been impacted by the pandemic, with consideration to both potential benefits (e.g., helping absent students stay connected to school) as well as harms. In the meantime, our findings suggest that education leaders will benefit from continuing to pay close attention to student attendance while also investing in more direct academic recovery efforts.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

[16] To avoid extrapolating too far beyond the results obtained in this paper, we do not attempt to generate a back-of-the-envelope estimate of the multi-year cumulative effects of absenteeism.

Ansari, A., & Gottfried, M. A. (2021). The grade-level and cumulative outcomes of absenteeism. *Child Development, 92*(4), e548–e564. https://doi.org/10.1111/cdev.13555

Aucejo, E. M., & Romano, T. F. (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review, 55*, 70–87. https://doi.org/10.1016/j.econedurev.2016.08.007

COVID-19 School Data Hub. (2022). *North Carolina state snapshot: Pandemic schooling mode and assessment outcome data series* [Data Brief 2022-14-NC-01]. https://www.covidschooldatahub.com

Curriculum Associates. (2023). *State of student learning in 2023.* https://www.curriculumassociates.com/research-and-efficacy/annual-report-the-state-of-student-learning-in-2023

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

Fahle, E., Kane, T. J., Patterson, T., Reardon, S. F., & Staiger, D. O. (2022). *Local achievement impacts of the pandemic.* Education Recovery Scorecard. https://educationrecoveryscorecard.org/wp-content/uploads/2022/10/Education-Recovery-Scorecard_Key-Findings_102822.pdf

Fahle, E., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2024). *The first year of pandemic recovery: A district-level analysis.* Education Recovery Scorecard. https://educationrecoveryscorecard.org/wp-content/uploads/2024/01/ERS-Report-Final-1.31.pdf

Fuller, S. C., Bastian, K. C., Swiderski, T., & Lauen, D. L. (2024). *Post-pandemic changes in absences and chronic absence rates in North Carolina: Updated through 2022-23.* Education Policy Initiative at Carolina. https://epic.unc.edu/wp-content/uploads/sites/1268/2024/03/Post-pandemic-Changes-in-Absences-and-Chronic-Absence-Rates-in-NC-2022-2023.pdf

Gershenson, S., Jacknowitz, A., & Brannegan, A. (2017). Are student absences worth the worry in U.S. primary schools? *Education Finance and Policy, 12*(2), 137–165. https://doi.org/10.1162/EDFP_a_00207

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

Goodman, J. (2014). *Flaking out: Student absences and snow days as disruptions of instructional time*. National Bureau of Economic Research. https://doi.org/10.3386/w20221

Gottfried, M. A. (2011). The detrimental effects of missing school: Evidence from urban siblings. *American Journal of Education, 117*(2), 147–182. https://doi.org/10.1086/657886

Gottfried, M. A. (2019). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education, 54*(1), 3–34. https://doi.org/10.1177/0042085915618709

Gottfried, M., & Ansari, A. (2022). Classrooms with high rates of absenteeism and individual success: Exploring students’ achievement, executive function, and socio-behavioral outcomes. *Early Childhood Research Quarterly, 59*(2), 215–227. https://doi.org/10.1016/j.ecresq.2021.11.008

Gottfried, M., Kim, P., & Fletcher, T. L. (2024). Do teachers perceive absent students differently? *AERA Open, 10*. https://doi.org/10.1177/23328584241259398

Gottfried, M. A., & Kirksey, J. J. (2017). "When" students miss school: The role of timing of absenteeism on students' test performance. *Educational Researcher, 46*(3), 119–130. https://doi.org/10.3102/0013189X17703945

Granados, A. (2021, December 14). On COVID, posting lesson plans, testing, race, and more: How this legislature dealt with education. *EdNC.* https://www.ednc.org/2021-12-14-a-legislative-recap-what-happened-this-long-session-of-the-general-assembly/

Jacob, B. A., & Stanjoevich, C. (2024). *Did COVID-19 shift the "grammar of schooling"?* EdWorkingPapers. https://doi.org/10.26300/8gy9-b516

Kuhfeld, M., & Lewis, K. (2024). *Recovery still elusive: 2023-24 student achievement highlights persistent achievement gaps and a long road ahead.* NWEA. https://www.nwea.org/research/publication/recovery-still-elusive-2023-24-student-achievement-highlights-persistent-achievement-gaps-and-a-long-road-ahead/

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

Lewis, K., & Kuhfeld, M. (2023). *Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery.* NWEA. https://www.nwea.org/research/publication/educations-long-covid-2022-23-achievement-data-reveal-stalled-progress-toward-pandemic-recovery/

Liu, J., Lee, M., & Gershenson, S. (2021). The short- and long-run impacts of secondary school absences. *Journal of Public Economics, 199*. https://doi.org/10.1016/j.jpubeco.2021.104441

Malkus, N. (2024). *Long COVID for public schools: Chronic absenteeism before and after the pandemic.* American Enterprise Institute. https://www.aei.org/wp-content/uploads/2024/01/Long-COVID-for-Public-Schools.pdf?x85095

Mervosh, S., & Paris, F. (2024, March 29). Why school absences have "exploded" almost everywhere. *The New York Times.* https://www.nytimes.com/interactive/2024/03/29/us/chronic-absences.html

Monk, D. H., & Ibrahim, M. A. (1984). Patterns of absence and pupil achievement. *American Educational Research Journal, 21*(2), 295–310.

National Center for Education Statistics [NCES]. (2022). *2022 NAEP Mathematics Assessment: Highlighted Results at Grades 4 and 8 for the Nation, States, and Districts* (NCES No. 2022124). U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022124

National Center for Education Statistics [NCES]. (2023). Rising absenteeism since onset of pandemic associated with NAEP 2022 score declines. *NAEP Plus.* U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/blog/attendance_and_naep_2022_score_declines.aspx

National Center for Education Statistics [NCES]. (2024). *2022–2023 Long-Term Trend Reading and Mathematics Assessment at Age 9 and 13* (NCES No. 2024038). U.S. Department of

Education, Institute of Education Sciences. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2024038

National Student Support Accelerator. (2023). *A snapshot of state tutoring policies.* https://studentsupportaccelerator.org/sites/default/files/Snapshot%20of%20State%20Tutoring%20Policies.pdf

Peters, S. J., Langi, M., Kuhfeld, M., & Lewis, K. (2023). *Unequal learning loss: How the COVID-19 pandemic influenced the academic growth of learners at the tails of the achievement distribution*. EdWorkingPapers. https://doi.org/10.26300/Z2EK-4937

Swiderski, T., Fuller, S. C., & Bastian, K. C. (2024, Sep. 9). Student-level attendance patterns show depth, breadth, and persistence of post-pandemic absenteeism. *Brookings Brown Center Chalkboard.* https://www.brookings.edu/articles/student-level-attendance-patterns-show-depth-breadth-and-persistence-of-post-pandemic-absenteeism/

The Learning Network. (2024, April 11). What students are saying about why school absences have 'exploded.' *The New York Times.* https://www.nytimes.com/2024/04/11/learning/what-students-are-saying-about-why-school-absences-have-exploded.html

Wei, W. (2024). Exploring patterns of absenteeism from prekindergarten through early elementary school and their associations with children's academic outcomes. *AERA Open, 10*. https://doi.org/10.1177/23328584241228212

Tables

Table 1

Descriptive statistics of the primary sample

	2016-17	2017-18	2018-19	2022-23
Math Std Score	0.313	0.320	0.312	0.138
	(0.839)	(0.841)	(0.854)	(0.910)
Days Absent	6.74	7.60	8.04	11.31
	(6.42)	(6.94)	(7.28)	(9.52)
Days Suspended	0.43	0.46	0.48	0.66
	(2.20)	(2.33)	(2.37)	(2.93)
Male	50.1%	50.0%	50.1%	50.9%
Black	20.4%	20.6%	20.8%	20.1%
Hispanic	17.0%	17.8%	18.2%	19.4%
White	53.8%	52.7%	51.6%	49.3%
Other Race/Ethnicity	8.7%	8.9%	9.4%	11.2%
English Learner	7.6%	11.3%	11.5%	11.2%
Student with Disability	6.2%	5.8%	5.7%	6.7%
AIG	18.9%	19.1%	18.9%	19.4%
Urban School	45.7%	46.2%	47.9%	47.7%
Change in Math Score	-0.059	-0.057	-0.060	-0.004
	(0.524)	(0.524)	(0.530)	(0.549)
Change in Days Absent	1.22	1.22	0.79	-0.62
	(5.42)	(5.97)	(6.01)	(8.45)
Change in Days Suspended	0.19	0.21	0.20	0.36
	(2.10)	(2.17)	(2.23)	(2.68)
N	313,345	328,639	347,248	313,660

Note. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 4 through 7. “AIG” = Academically or Intellectually Gifted. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Days Absent and Days Suspended are normalized to a 180-day school year by multiplying a student’s percent of days absent or suspended by 180. Variables labeled “Change in” refer to the average difference from the prior year.

Table 2

Mean changes in math standardized scores and days absent, 2018-19 vs. 2022-23

	Math Std Score		Days Absent		Peer Absences	
	(1)	(2)	(3)	(4)	(5)	(6)
Analytic Sample						
2022-23	-0.174***	-0.176***	3.27***	3.27***	3.64***	3.65***
	(0.0154)	(0.0106)	(0.083)	(0.077)	(0.095)	(0.087)
Covs?	N	Y	N	Y	N	Y
N	660,908	660,908	660,908	660,908	660,908	660,908
All Students						
2022-23	-0.166***	-0.163***	3.81***	3.82***	3.91***	3.90***
	(0.0178)	(0.0109)	(0.102)	(0.091)	(0.106)	(0.095)
Covs?	N	Y	N	Y	N	Y
N	890,456	890,456	890,456	890,456	890,456	890,456

Note. "All students" includes students with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 4 through 7. "Analytic sample" further restricts to students in the top 80% of the achievement distribution in the prior year. Standard errors, clustered by school-year, shown in parentheses. Covariates in even-numbered models include student race/ethnicity, sex, English Learner status, disability status, academically or intellectually gifted status, and an indicator for being in an urban school. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Days Absent is normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Peer Absences refers to the average number of (normalized) days absent among peers in the focal student's math class.

Table 3

Relationship between absenteeism and achievement, pre- and post-COVID, student first-differences models

	’16 to ’17 (1)	’17 to ’18 (2)	’18 to ’19 (3)	’22 to ’23 (4)	’16 to ’17 (5)	’17 to ’18 (6)	’18 to ’19 (7)	’22 to ’23 (8)
Days Abs (x10)	-0.063***	-0.063***	-0.063***	-0.040***	-0.054***	-0.054***	-0.055***	-0.032***
	(0.0028)	(0.0022)	(0.0021)	(0.0018)	(0.0027)	(0.0022)	(0.0020)	(0.0018)
Days Susp (x10)					-0.134***	-0.139***	-0.126***	-0.140***
					(0.0065)	(0.0060)	(0.0058)	(0.0048)
N per year	313,345	328,639	347,248	313,660	313,345	328,639	347,248	313,660

Note. Standard errors, in parentheses, clustered by school-year. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2015-16 to 2016-17. Sample includes students in the top 80% of the achievement distribution in the model’s baseline year, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in the model’s baseline year and who were observed in the next grade level in the second year.
* p < .05, ** p < .01, *** p < .001

Table 4

Relationship between absenteeism and achievement, pre- and post-COVID, student first-differences models, peer effects models

	’17 to ’18	’18 to ’19	’22 to ’23	’17 to ’18	’18 to ’19	’22 to ’23
	(1)	(2)	(3)	(4)	(5)	(6)
Days Abs (x10)	-0.064***	-0.064***	-0.039***	-0.055***	-0.055***	-0.030***
	(0.0021)	(0.0021)	(0.0018)	(0.0021)	(0.0020)	(0.0018)
Peer Abs (x10)	-0.020*	-0.038***	0.001	-0.011	-0.035***	0.004
	(0.0097)	(0.0086)	(0.0061)	(0.0100)	(0.0094)	(0.0066)
Days Susp (x10)				-0.113***	-0.102***	-0.115***
				(0.0060)	(0.0062)	(0.0048)
Peer % Suspended				-0.264***	-0.242***	-0.289***
				(0.0217)	(0.0213)	(0.0209)
Peer Prior Math Avg				-0.047***	-0.053***	-0.056***
				(0.0048)	(0.0047)	(0.0048)
N per year	233,189	246,639	211,170	233,189	246,639	211,170

Note. Standard errors, in parentheses, clustered by school-year. Coefficients and standard errors of variables with “(x10)” show estimates of a 10-unit increase in the variable. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2016-17 to 2017-18. Sample includes students in the top 80% of the achievement distribution in the model’s baseline year, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 4 through 6 in the model’s baseline year and who were observed in the next grade level in the second year.

* p < .05, ** p < .01, *** p < .001

Table 5

Relationship between absenteeism and achievement, pre- and post-COVID, student FD models, racial/ethnic subgroups

	2018 to 2019 (1)	2022 to 2023 (2)
Black		
Days Abs (x10)	-0.050***	-0.033***
	(0.0036)	(0.0027)
N per year	72,237	63,109
White		
Days Abs (x10)	-0.055***	-0.039***
	(0.0027)	(0.0024)
N per year	179,169	154,540
Hispanic		
Days Abs (x10)	-0.057***	-0.032***
	(0.0041)	(0.0030)
N per year	63,088	60,850
Covariates?	Y	Y

Note. Standard errors, in parentheses, clustered by school-year. Coefficients and standard errors of variables with "(x10)" show estimates of a 10-unit increase in the variable. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2017-18 to 2018-19. Sample includes students in the top 80% of the achievement distribution in the model's baseline year, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in the model's baseline year and who were observed in the next grade level in the second year. Covariates (not shown) include days suspended.

* p < .05, ** p < .01, *** p < .001

Table 6

Main results by baseline achievement subgroups

	2018 to 2019 (1)	2022 to 2023 (2)
Quintile 1 (Lowest) in Base Year	-0.023***	-0.016***
	(0.0023)	(0.0016)
N	81,450	104,486
Quintile 2 in Base Year	-0.053***	-0.028***
	(0.0030)	(0.0024)
N	94,070	86,382
Quintile 3 in Base Year	-0.063***	-0.032***
	(0.0033)	(0.0031)
N	87,305	74,473
Quintile 4 in Base Year	-0.058***	-0.024***
	(0.0037)	(0.0035)
N	87,453	67,183
Quintile 5 (Highest) in Base Year	-0.048***	-0.034***
	(0.0042)	(0.0038)
N	78,420	59,880
Covariates?	Y	Y

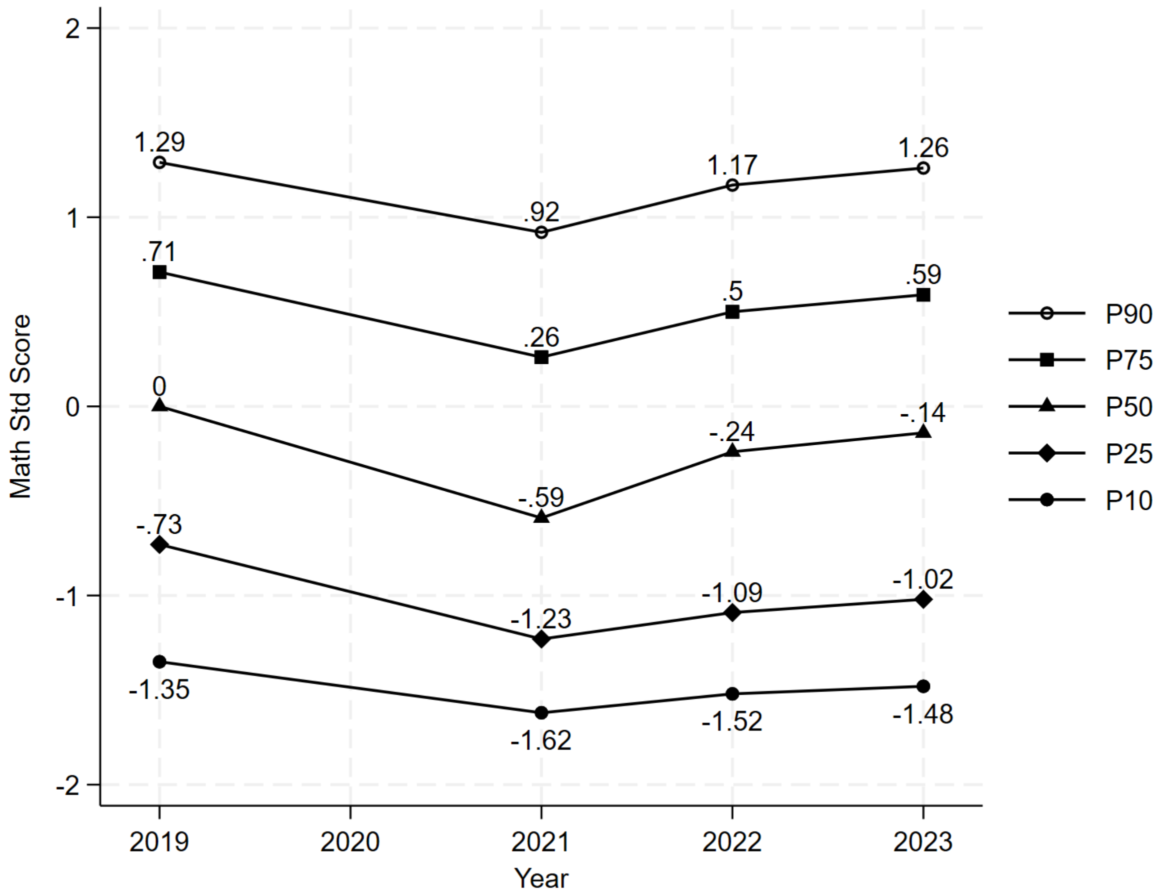
Note. Standard errors, in parentheses, clustered by school-year. Coefficients and standard errors of variables with "(x10)" show estimates of a 10-unit increase in the variable. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2017-18 to 2018-19. Sample includes students with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in the model's baseline year and who were observed in the next grade level in the second year. For Model 1, "Quintile 1" refers to students whose standardized math score in 2017-18 was in the bottom quintile of their cohort. For Model 2, "Quintile 1" refers to students whose scale math score in 2021-22 would have put them in the baseline quintile of the 2017-18 cohort. Covariates (not shown) include days suspended.

* p < .05, ** p < .01, *** p < .001

Figures

Figure 1

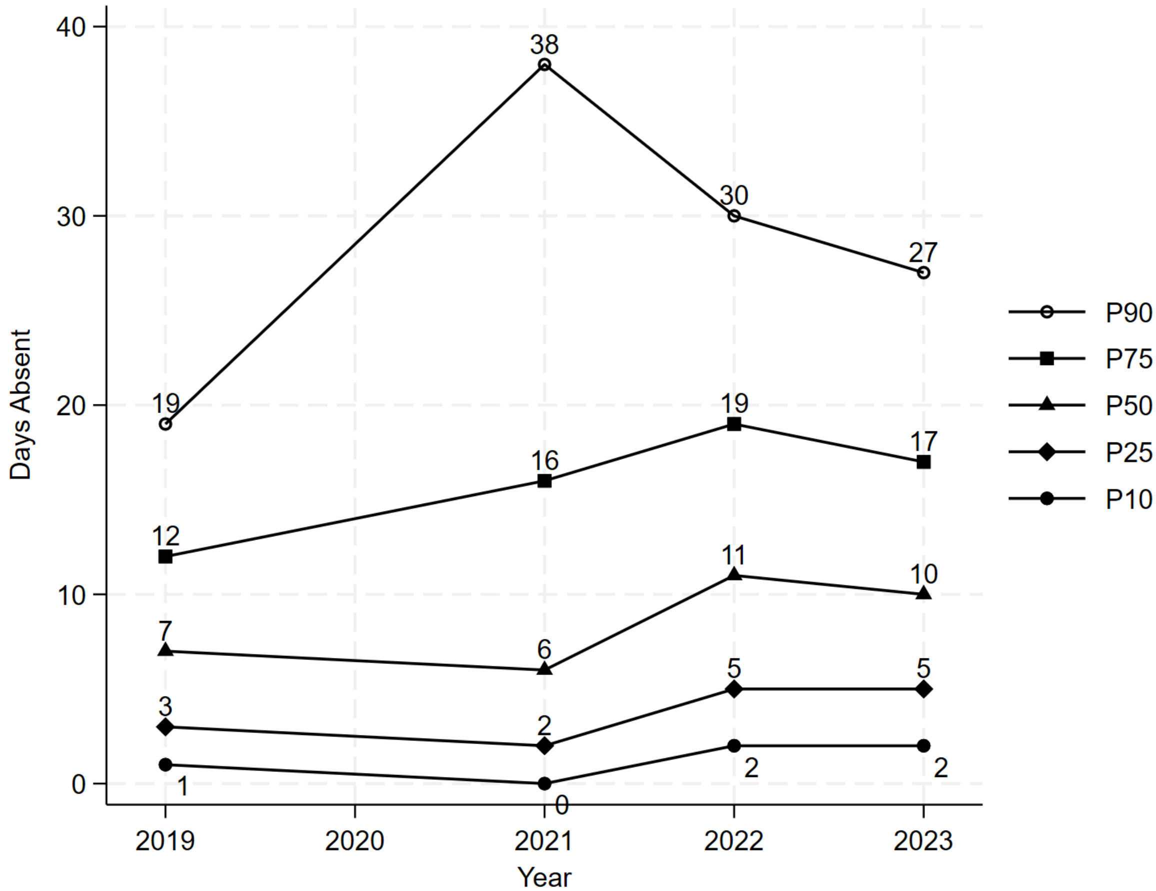
Math test score quantiles over time



Note. Figure shows the 10th, 25th, 50th, 75th, and 90th percentile of math standardized test scores in 2018-19, 2020-21, 2021-22, and 2022-23. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Sample includes all students in NC public schools in grades 3 through 7 with a valid end-of-grade math exam score.

Figure 2

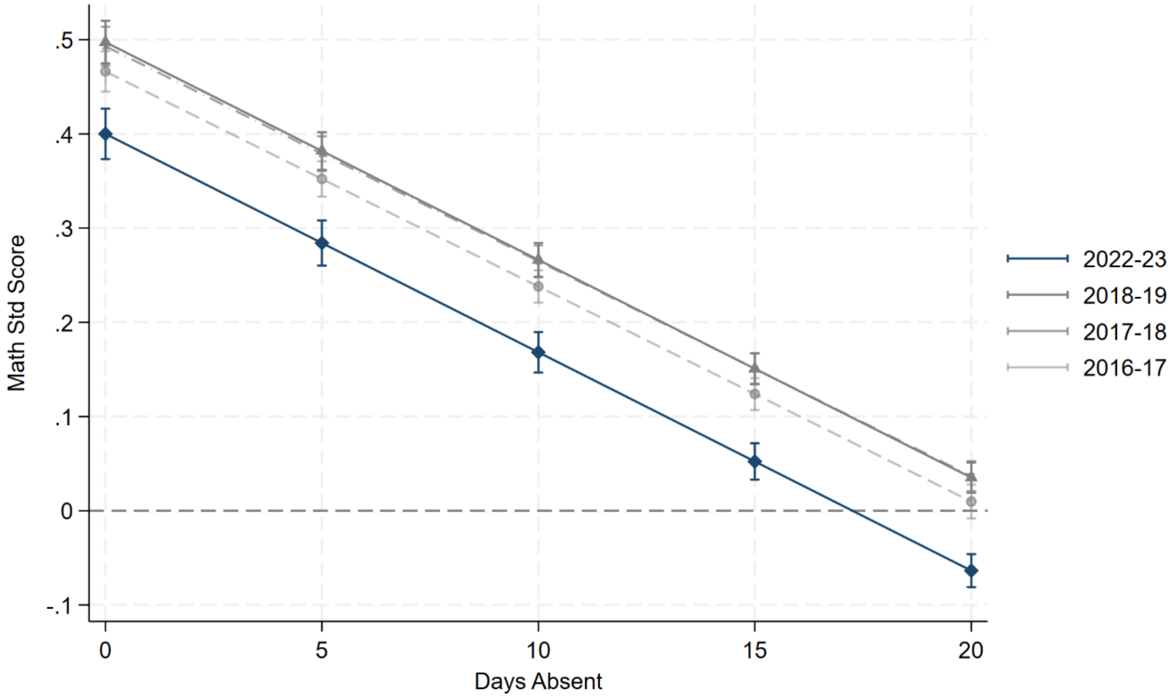
Absence quantiles over time



Note. Figure shows the 10th, 25th, 50th, 75th, and 90th percentile of math standardized test scores in 2018-19, 2020-21, 2021-22, and 2022-23. Days Absent is normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Sample includes all students in NC public schools in grades 3 through 7 with valid attendance data.

Figure 3

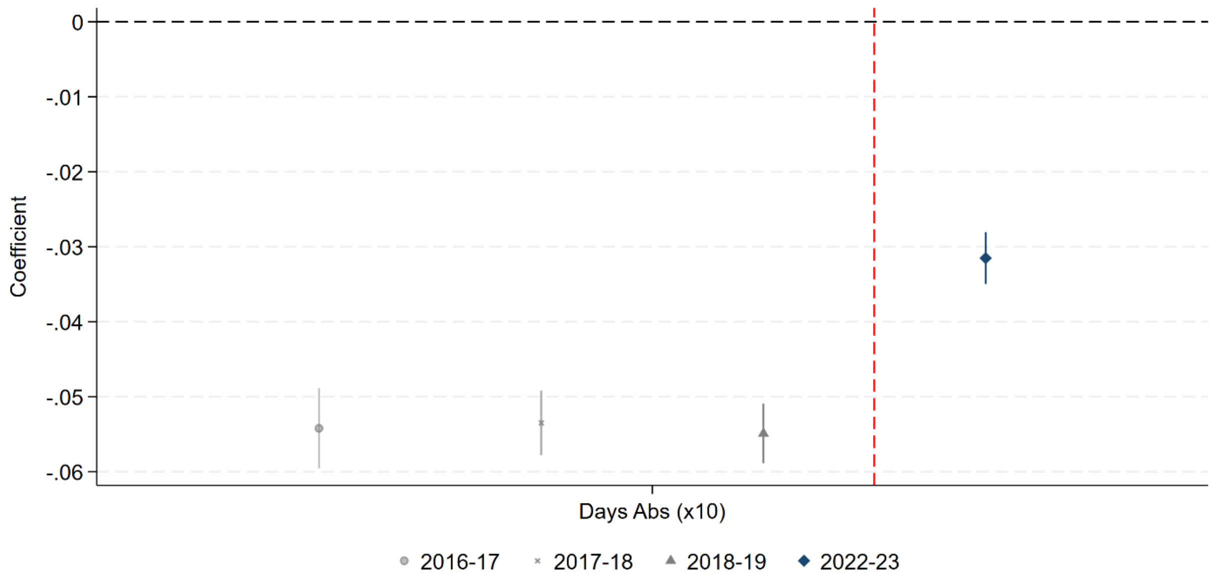
Simple correlation between math scores and days absent in each school year



Note. This figure depicts marginal predicted values at select values of days absent, obtained from simple OLS regressions of math scores on days absent in each of the four years shown (as shown in Appendix Table A4). Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Days Absent is normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Sample includes students who were in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled in the model's focal year, among students in NC public schools who were in grades 4 through 7 in the model's focal year and who were observed in the prior grade level in the previous year.

Figure 4

Coefficient plot of first-differences estimates of effect of absenteeism on achievement



Note. Figure depicts coefficients from Table 3, Models 5 through 8 ("2016-17" refers to Model 5, which included data from 2015-16 to 2016-17). Confidence intervals, clustered by school-year, are shown via lines extending from points.

Online Appendix Tables

Appendix Table A1

Additional descriptive statistics on math test scores and absences

	P5	P10	P25	P50	P75	P90	P95
Change in Days Abs							
2016-17	-6	-4	-1	1	4	7	10
2017-18	-7	-5	-2	1	4	8	11
2018-19	-8	-5	-2	0	3	7	10
2022-23	-14	-10	-5	0	4	8	12
Change in Std Math Score							
2016-17	-0.94	-0.74	-0.39	-0.06	0.30	0.60	0.78
2017-18	-0.93	-0.72	-0.41	-0.04	0.28	0.60	0.78
2018-19	-0.94	-0.73	-0.41	-0.05	0.29	0.62	0.80
2022-23	-0.90	-0.70	-0.37	0.01	0.37	0.70	0.90

Note. This table shows the distribution of changes in days absent or math scores from the prior year. "P5" refers to the 5th percentile. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 4 through 7. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score/percentile rank that they would have earned in 2018-19). Days Absent are normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Sample sizes: 2016-17 = 313,345; 2017-18 = 328,639; 2018-19 = 347,248; 2022-23 = 313,660.

Appendix Table A2

Descriptive statistics of characteristics of math classroom peers

	2017-18	2018-19	2022-23
Peer Avg Days Absent	8.32	8.75	12.40
	(2.84)	(3.00)	(4.24)
% of Peers Suspended	13.7%	14.0%	17.0%
Peer Avg Prior Math Score	0.165	0.167	-0.051
	(0.622)	(0.631)	(0.649)
Change in Peer Avg Days Absent	1.41	0.77	-0.94
	(3.04)	(2.97)	(4.71)
Change in % of Peers Suspended	0.042	0.036	0.059
	(0.134)	(0.135)	(0.147)
Change in Peer Avg Prior Math Score	0.038	0.040	0.264
	(0.527)	(0.534)	(0.539)
N	233,189	246,639	211,170

Note. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 5 through 7. Standardized math scores post-pandemic are anchored to the 2018-19 distribution (i.e., scale scores after 2019-20 are assigned the standardized score/percentile rank that they would have earned in 2018-19). Days Absent is normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Variables labeled "Change in" refer to the average difference from the prior year.

Appendix Table A3

Additional descriptive statistics on classroom peer absenteeism

	P5	P10	P25	P50	P75	P90	P95
Change in Peer Avg Days Abs							
2017-18	-3.0	-2.0	-0.5	1.1	3.1	5.3	6.8
2018-19	-3.8	-2.7	-1.0	0.6	2.4	4.4	5.8
2022-23	-8.7	-6.7	-3.7	-0.8	1.9	4.6	6.4

Note. This table shows the distribution of changes in the absence rate of a student's classroom peers from the prior year. "P5" refers to the 5th percentile. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 5 through 7. Days Absent are normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Sample sizes: 2017-18 = 233,189; 2018-19 = 246,639; 2022-23 = 211,170.

Appendix Table A4

Correlations between current-year absences and current-year test scores

	2017	2018	2019	2023
	(1)	(2)	(3)	(4)
Days Abs (x10)	-0.023***	-0.023***	-0.023***	-0.023***
	(0.0005)	(0.0004)	(0.0004)	(0.0004)
N	313,345	328,639	347,248	313,660

Note. Standard errors, in parentheses, clustered by school-year. Column headers refer to the spring of the school year included in the model (e.g., Model 1 shows results for school year 2016-17). Coefficients and standard errors of show estimates of a 10-unit increase in days absent. Sample includes students who were in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled in the model's focal year, among students in NC public schools who were in grades 4 through 7 in the model's focal year and who were observed in the prior grade level in the previous year.
* p < .05, ** p < .01, *** p < .001

Appendix Table A5

Sensitivity to sample specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2018 to 2019								
Days Abs (x10)	-0.055***	-0.052***	-0.049***	-0.055***	-0.052***	-0.056***	-0.044***	-0.066***
	(0.0020)	(0.0017)	(0.0034)	(0.0020)	(0.0021)	(0.0021)	(0.0016)	(0.0014)
Days Abs^2 (x10)			-0.0019*					
			(0.0008)					
Days Susp (x10)	-0.126***	-0.110***	-0.126***	-0.126***	-0.122***	-0.130***	-0.054***	-0.142***
	(0.0058)	(0.0054)	(0.0058)	(0.0058)	(0.0060)	(0.0061)	(0.0038)	(0.0054)
N	347,248	347,248	347,248	347,248	342,971	319,863	428,698	347,248
2022 to 2023								
Days Abs (x10)	-0.032***	-0.033***	-0.029***	-0.031***	-0.029***	-0.033***	-0.033***	-0.049***
	(0.0018)	(0.0014)	(0.0033)	(0.0017)	(0.0019)	(0.0018)	(0.0014)	(0.0012)
Days Abs^2 (x10)			-0.0006					
			(0.0005)					
Days Susp (x10)	-0.140***	-0.105***	-0.139***	-0.134***	-0.142***	-0.142***	-0.052***	-0.117***
	(0.0048)	(0.0043)	(0.0048)	(0.0047)	(0.0050)	(0.0050)	(0.0036)	(0.0045)
N	313,660	313,660	313,660	313,660	306,079	281,895	392,404	313,660

Note. Coefficients and standard errors of show estimates of a 10-unit increase in days absent or suspended. Models: (1) Main sample estimate; (2) Additional covariates for English Learner status, disability status, academically or intellectually gifted status, and whether the student was in an urban school, and a school fixed effect; (3) Main model, plus Days Abs Squared; (4) Main model, outcome = standardized score rather than anchored standardized score for post-pandemic; (5) Main model, omit students with outlier changes in absences or test scores; (6) Main model, omit charter school students; (7) Main model, include all students, including the bottom baseline achievement quintile; (8) Value-added model, with controls for student grade-level, race/ethnicity, sex, economic disadvantage status, English Learner status, disability status, days suspended, and an indicator of having attended multiple schools; and a school fixed effect. All standard errors are clustered by school-year.

Appendix Table A6

Sensitivity to sample specifications, peer effects models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2018 to 2019								
Days Abs (x10)	-0.055***	-0.054***	-0.050***	-0.055***	-0.051***	-0.056***	-0.043***	-0.063***
	(0.0020)	(0.0019)	(0.0035)	(0.0020)	(0.0022)	(0.0021)	(0.0017)	(0.0016)
Days Abs^2 (x10)			-0.0015					
			(0.0008)					
Peer Abs (x10)	-0.035***	-0.016*	0.054*	-0.035***	-0.034***	-0.038***	-0.013	-0.025*
	(0.0094)	(0.0069)	(0.0270)	(0.0094)	(0.0090)	(0.0100)	(0.0075)	(0.0104)
Peer Abs^2 (x10)			-0.046***					
			(0.0121)					
N	246,639	246,639	246,639	246,639	243,698	227,455	303,537	246,611
2022 to 2023								
Days Abs (x10)	-0.030***	-0.031***	-0.030***	-0.030***	-0.028***	-0.031***	-0.031***	-0.046***
	(0.0018)	(0.0016)	(0.0032)	(0.0017)	(0.0020)	(0.0019)	(0.0015)	(0.0014)
Days Abs^2 (x10)			-0.000					
			(0.0006)					
Peer Abs (x10)	0.004	0.011	0.012	0.006	-0.002	0.002	-0.000	-0.039***
	(0.0066)	(0.0054)	(0.0201)	(0.0065)	(0.0065)	(0.0072)	(0.0055)	(0.0084)
Peer Abs^2 (x10)			-0.003					
			(0.0059)					
N	211,170	211,170	211,170	211,170	204,883	189,574	262,078	211,122

Note. Coefficients and standard errors of show estimates of a 10-unit increase in own/peer days absent or suspended. Models: (1) Main sample estimate; (2) Additional covariates for English Learner status, disability status, academically or intellectually gifted status, and whether the student was in an urban school, and a school fixed effect; (3) Main model, plus Days Abs Squared; (4) Main model, outcome = standardized score rather than anchored standardized score for post-pandemic; (5) Main model, omit students with outlier changes in absences or test scores; (6) Main model, omit charter school students; (7) Main model, include all students, including the bottom baseline achievement quintile; (8) Value-added model, with controls for student grade-level, race/ethnicity, sex, economic disadvantage status, English Learner status, disability status, days suspended, and an indicator of having attended multiple schools; math classroom peer-level prior math scores, suspensions, sex, race/ethnicity, economic disadvantage status, English Learner status, disability status; the number of students in the math class; and a school fixed effect. All standard-errors are clustered by school-year.

Appendix Table A7

Relationship between absenteeism and achievement, 2020-21 to 2021-22, student first-differences models and value-added models

	First-Difference		VAM
	(1)	(2)	(3)
Days Abs (x10)	-0.003*	-0.002	-0.047***
	(0.0014)	(0.0014)	(0.0012)
Days Susp (x10)		-0.178***	-0.137***
		(0.0059)	(0.0051)
N per year	291,879	291,879	291,879

Note. Coefficients and standard errors of show estimates of a 10-unit increase in days absent or suspended. Standard errors, in parentheses, clustered by school-year. Model 3 ("VAM") refers to a value-added model, including controls (not shown) for student grade-level, race/ethnicity, sex, economic disadvantage status, English Learner status, disability status, days suspended, and an indicator of having attended multiple schools; and a school fixed effect. Sample includes students in the top 80% of the achievement distribution in 2020-21, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in 2020-21 and who were observed in the next grade level in the second year.

* p < .05, ** p < .01, *** p < .001

Appendix Table A8

Main results by other student demographic subgroups

	2018 to 2019 (1)	2022 to 2023 (2)
Gr 3 to 4	-0.050***	-0.041***
	(0.0041)	(0.0034)
N	88,020	79,073
Gr 4 to 5	-0.053***	-0.029***
	(0.0038)	(0.0033)
N	87,722	79,281
Gr 5 to 6	-0.064***	-0.034***
	(0.0043)	(0.0038)
N	88,137	77,054
Gr 6 to 7	-0.051***	-0.022***
	(0.0034)	(0.0031)
N	83,369	78,252
Male	-0.055***	-0.030***
	(0.0026)	(0.0021)
N	174,106	159,500
Female	-0.054***	-0.033***
	(0.0026)	(0.0022)
N	173,142	154,160
SWD	-0.058***	-0.027***
	(0.0060)	(0.0044)
N	19,901	21,104
Not SWD	-0.054***	-0.032***
	(0.0021)	(0.0018)
N	327,347	292,556
Covs?	Y	Y

Note. Coefficients and standard errors of show estimates of a 10-unit increase in days absent. Standard errors, in parentheses, clustered by school-year. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2017-18 to 2018-19. Sample includes students in the top 80% of the achievement distribution in the model's baseline year, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in the model's baseline year and who were observed in the next grade level in the second year. Covariates (not shown) include days suspended. SWD = "Student with Disability." "Gr 3 to 4" refers to students who were in Grade 3 in the baseline year and Grade 4 in the next year.

* p < .05, ** p < .01, *** p < .001

Appendix Table A9

Main results by student demographic subgroups, peer effects models

		2018 to 2019 (1)	2022 to 2023 (2)
Black	Days Abs (x10)	-0.050***	-0.030***
		(0.0039)	(0.0032)
	Peer Abs (x10)	-0.021	0.009
		(0.012)	(0.0080)
N		50,548	41,266
White	Days Abs (x10)	-0.053***	-0.037***
		(0.0028)	(0.0024)
	Peer Abs (x10)	-0.037**	0.001
		(0.012)	(0.0088)
N		128,316	105,570
Hispanic	Days Abs (x10)	-0.056***	-0.028***
		(0.0044)	(0.0034)
	Peer Abs (x10)	-0.017	-0.020*
		(0.013)	(0.0088)
N		45,163	41,921
Gr 4 to 5	Days Abs (x10)	-0.054***	-0.030***
		(0.0037)	(0.0031)
	Peer Abs (x10)	-0.020	0.006
		(0.0144)	(0.0111)
N		83,556	71,422
Gr 5 to 6	Days Abs (x10)	-0.058***	-0.036***
		(0.0035)	(0.0031)
	Peer Abs (x10)	-0.042*	0.003
		(0.017)	(0.0122)
N		83,956	69,518
Gr 6 to 7	Days Abs (x10)	-0.050***	-0.024***
		(0.0032)	(0.0028)
	Peer Abs (x10)	-0.041**	0.012
		(0.0133)	(0.0095)
N		79,099	70,172
Covs?		Y	Y

Note. Coefficients and standard errors of show estimates of a 10-unit increase in own/peer days absent. Standard errors, in parentheses, clustered by school-year. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2017-18 to 2018-19. Sample includes students in the top 80% of the achievement distribution in the model's baseline year, with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 4 through 6 in the model's baseline year and who were observed in the next grade level in the second year. Covariates (not shown) include days suspended, classroom peer suspension rate, and classroom peer average prior math score. "Gr 4 to 5" refers to students who were in Grade 4 in the baseline year and Grade 5 in the next year.
* p < .05, ** p < .01, *** p < .001

Appendix Table A10

Main results by baseline achievement subgroups, based on position in non-anchored distribution

	2018 to 2019 (1)	2022 to 2023 (2)
Lowest-achieving in Base Year	-0.023***	-0.012***
	(0.0023)	(0.0017)
N	81,450	78,744
Ach Quantile 2 in Base Year	-0.053***	-0.024***
	(0.0030)	(0.0023)
N	94,070	76,477
Ach Quantile 3 in Base Year	-0.063***	-0.029***
	(0.0033)	(0.0028)
N	87,305	80,247
Ach Quantile 4 in Base Year	-0.058***	-0.028***
	(0.0037)	(0.0031)
N	87,453	84,058
Highest-achieving in Base Year	-0.048***	-0.032***
	(0.0042)	(0.0036)
N	78,420	72,878
Covs?	Y	Y

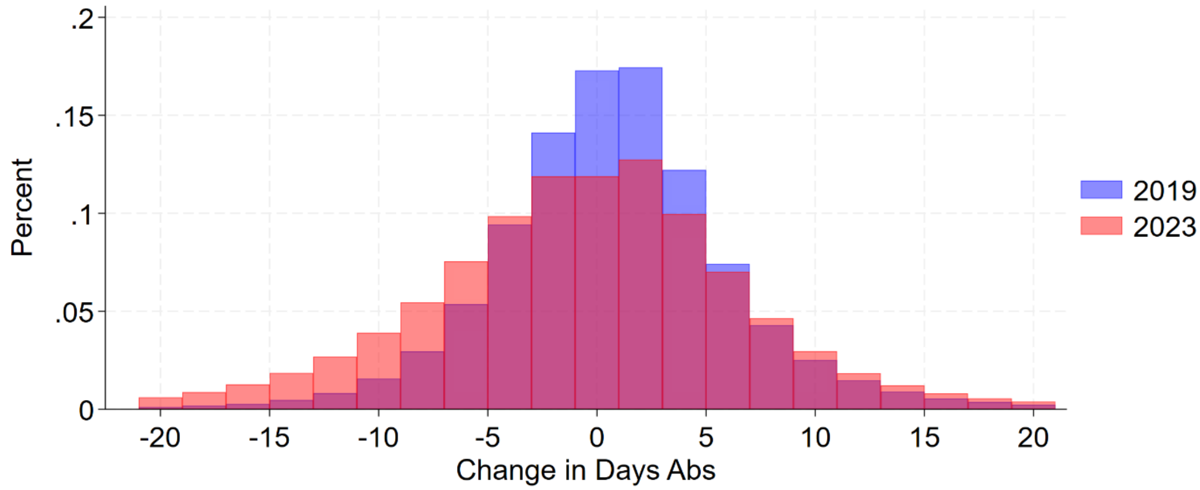
Note. Coefficients and standard errors of show estimates of a 10-unit increase in days absent. Standard errors, in parentheses, clustered by school-year. Models are restricted to the school years indicated in their column header. For example, Model 1 is restricted to school years 2017-18 to 2018-19. Sample includes students with <50% of days absent or suspended and at least 90 days enrolled in each year, among students in NC public schools who were in grades 3 through 6 in the model's baseline year and who were observed in the next grade level in the second year. "Quintile 1" refers to students whose standardized math score in the baseline year of the model was in the bottom quintile of their cohort. Covariates (not shown) include days suspended.

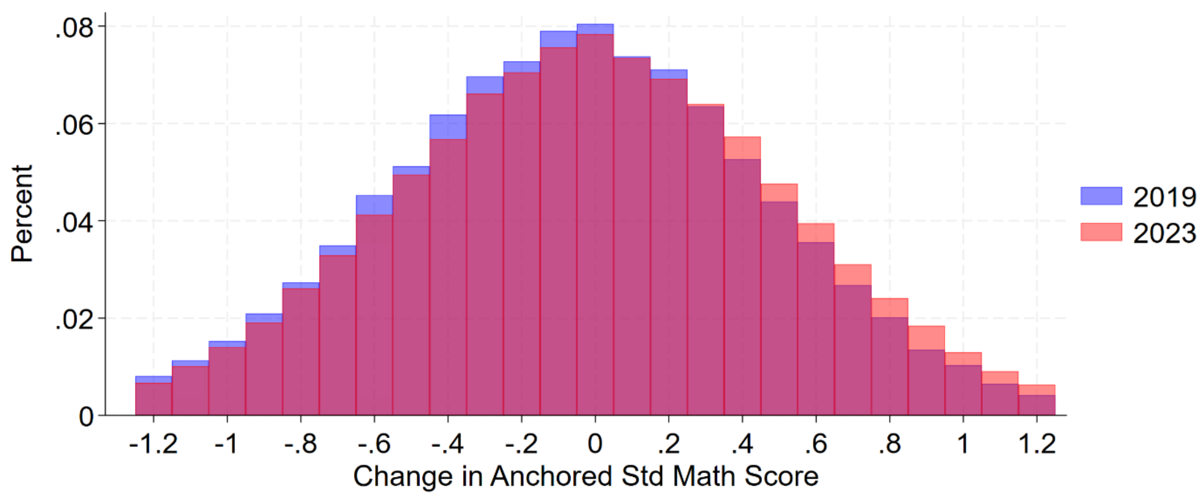
* p < .05, ** p < .01, *** p < .001

Online Appendix Figures

Appendix Figure A1

Histogram of Change in Days Absent and Change in Math Scores, 2018-19 vs. 2022-23

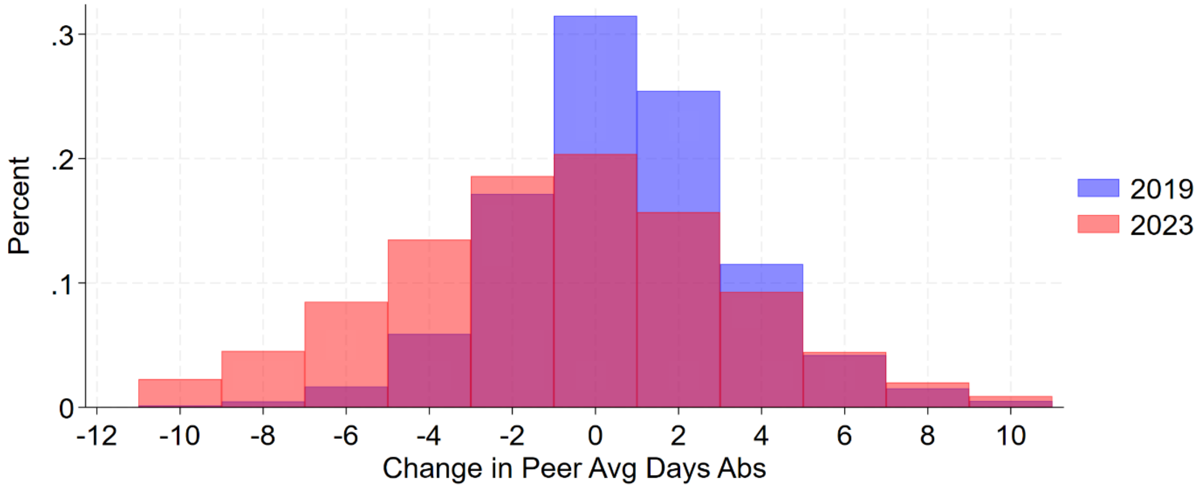




Note. Figure depicts the distribution of changes in days absent (top) and changes in standardized math scores (bottom) across students in 2018-19 (blue) and 2022-23 (red) relative to the prior year. Charts are restricted to values between approximately the 1st and 99th percentile of the combined distribution to increase readability and protect confidentiality. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Days Absent is normalized to a 180-day school year by multiplying a student's percent of days absent by 180. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 4 through 7.

Appendix Figure A2

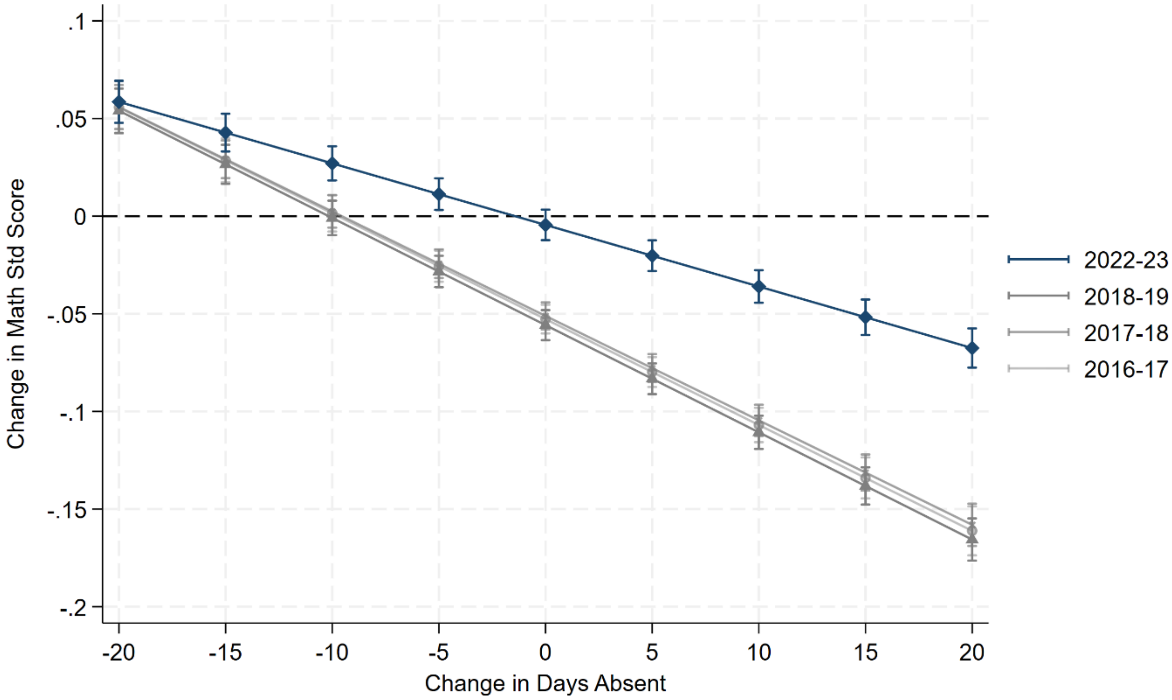
Histogram of Change in Peer Avg Days Absent, 2018-19 vs. 2022-23



Note. Figure depicts the distribution of changes in the absence rate of classroom peers across students in 2018-19 (blue) and 2022-23 (red) relative to the prior year. Chart is restricted to values between approximately the 1st and 99th percentile of the combined distribution to increase readability and protect confidentiality. Sample includes students in the top 80% of the achievement distribution in the prior year, with <50% of days absent or suspended and at least 90 days enrolled, among students in NC public schools in grades 5 through 7.

Appendix Figure A3

Margins plot: Predicted test scores by days absent and year



Note. Figure depicts marginal predicted values of student standardized math scores at selected values of changes in days absent, obtained from first-difference models with controls for days suspended, in school years 2016-17, 2017-18, 2018-19, and 2022-23 (see Table 3, Models 5 through 8), where "change in days absent" refers to the change from the prior year. Standardized math scores in 2022-23 are anchored to the 2018-19 distribution (i.e., scale scores in 2022-23 are assigned the standardized score that they would have earned in 2018-19). Days Absent and Days Suspended are normalized to a 180-day school year by multiplying a student's percent of days absent or suspended by 180. Sample includes students in the top 80% of the achievement distribution in the model's baseline year (i.e., the year prior to the year indicated in the legend), with <50% of days absent or suspended and at least 90 days enrolled in each year used in the model, among students in NC public schools who were in grades 3 through 6 in the model's baseline year and who were observed in the next grade level in the second year.