The Path of Student Learning Delay During the COVID-19 Pandemic: Evidence from Michigan

June, 2025

Katharine O. Strunk
kstrunk@upenn.edu
University of Pennsylvania

Tara Kilbride
kilbrid9@msu.edu
Michigan State University & EPIC

Scott Imberman
imberman@msu.edu
Michigan State University & NBER

Seth Walker
walke893@msu.edu
Michigan State University

Dongming Yu
yudongmi@msu.edu
Michigan State University

Bryant Hopkins
bryant.hopkins@bateswwhite
Bates White Consulting

Abstract: Educators and policymakers have been concerned that the COVID-19 pandemic led to substantial delays in learning due to disruptions, anxiety, and remote schooling. We study student achievement patterns over the pandemic using a combination of state summative and higher frequency benchmark assessments for 5th and 6th grade students in Michigan. Comparing pre-pandemic to post-pandemic cohorts we find that math achievement growth dropped by 0.20 standard deviations more than expected, between 2019 and 2022. These drops were larger for Black, Latino, and economically disadvantaged students, as well as students in districts that were at least partially remote in 2021-22. Benchmark assessment results are consistent with summative assessments and show sharp drops in 2020-21 with a partial recovery in math. Reductions in ELA growth were generally small and statistically insignificant. Both assessments and subjects, however, show recovery stalling at a level below pre-pandemic achievement.

Acknowledgements: We would like to thank the State of Michigan and private philanthropy organizations who financially supported this work. We would also like to thank the MERI-Michigan Education Data Center (MEDC) who provided the data used in this work. MEDC data is modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI). Results, information and opinions solely represent the analysis, information and opinions of the author(s) and are not endorsed by, or reflect the views or positions of, grantors, MDE and CEPI or any employee thereof.

1. Introduction

The COVID-19 pandemic severely impacted student achievement across the United States. Nationally, average test scores in fall 2021 were substantially below historic averages and academic recovery since then has been slow (Goldhaber et al., 2022; Kuhfeld & Lewis, 2022). For example, spring 2022 end-of-year testing outcomes from multiple states show that student achievement continued to trail pre-pandemic levels (e.g., Halloran et al., 2023; Kogan, 2022; Idaho State Department of Education, 2022; Tennessee Department of Education, 2022; Texas Education Agency, 2022; Sass & Ali, 2022) even after the initial phases of the pandemic. Similarly, results from the 2022 administration of the National Assessment of Education Progress (NAEP) show historically large decreases in average achievement and widening gaps between high- and low-performing students between 2019 and 2022 (National Center for Education Statistics, 2022). Pandemic impacts have been particularly acute for certain student subgroups, including students of color and those receiving additional services, as well as students attending high-poverty schools and elementary schools, those who learned remotely, and those with lower baseline achievement (e.g., Goldhaber et al., 2023; Kilbride et al., 2022). Recent work has indicated that pandemic recovery may have stalled (Lewis & Kuhfeld, 2023) and NAEP results show that disparities in student achievement continued to widen between 2022 and 2024 (National Center for Education Statistics, 2025). Considering these findings, it is imperative that research continues to document achievement trends so that educators, policymakers, and the public can better understand how the pandemic and associated school disruptions affected and continue to affect students' academic development.

This paper uses student achievement measures from Michigan's summative end-of-year tests (the Michigan Student Test of Educational Progress, M-STEP) and fall and spring NWEA

MAP Growth and Curriculum Associates' i-Ready benchmark assessments to assess achievement growth and trajectories during the pandemic. A particularly useful benefit of combining these two data sources is that we can examine both the total change in achievement during the pandemic through spring 2023, including the recovery, as well as how achievement progressed *within* the pandemic-affected school years. The relatively high frequency of our data allows us to show how the pandemic recovery was not constant, but rather involved a sharp drop, followed by partial recovery, and subsequent stall-out. This is key information for policymakers if and when they need to navigate a potential future mass disruption to schooling such as a natural disaster or flu pandemic. We also examine heterogeneity in performance across students with different demographic characteristics and those who participated in different modes of instruction (e.g., fully in-person, fully remote, or hybrid instruction).

While some work has looked at the impacts of COVID on achievement using nationwide data (Goldhaber et al, 2023), our work makes several key contributions on top of the existing literature. First, by combining data on state summative assessments and benchmark assessments we show that the patterns in these two types of tests are quite similar, both overall and across different types of students. Second, with data from both assessments that now goes through the spring of 2023, we establish that the overall recovery has stalled, but not for everyone. For example, relative to their pre-pandemic starting points, Black students continue to make larger gains than White students to the point where the former has erased the disproportionate losses from the initial pandemic stages. Further, we provide a deep-dive analysis into the dynamics of student achievement during the pandemic looking across race, economic status, and instructional modality. Most other studies either focus on just one of these heterogeneous factors or consider

shorter-term trends. This study brings these multiple factors together into a single narrative of how the pandemic influenced education and how it continues to do so.

We first compare three-year growth M-STEP outcomes for a "pre-pandemic cohort" that includes students who completed either the math or ELA assessment in both 2016 and 2019, and students in a "pandemic cohort" that completed the M-STEP in 2019 and 2022.[1] Our results show that growth in math was 0.20 standard deviations (SD) lower in the pandemic cohort, but there was only a statistically insignificant reduction in ELA achievement growth of 0.03 SD. Math growth differed substantially by student type with Black, Latino/a, economically disadvantaged, and students who spent more time in remote instruction all showing growth between 0.04 and 0.07 SD lower than the relevant comparison. There were no significant differences across any of these groups for ELA, however.

We also examine trends in achievement on nationally normed benchmark assessments across the fall 2020 through spring 2023 testing periods and M-STEP from spring 2019 through spring 2023. These analyses provide additional insight into students' achievement trajectories by capturing more granular changes during the school years that were directly impacted by the pandemic. To align our samples and ensure students have benchmark and M-STEP exams throughout the pandemic period, we focus on students in grades 5 and 6 at the start of the pandemic.[2] In particular, Michigan students were scoring much farther behind national norms in

[1] M-STEP exams go from grades 3 through 7 with 8th grade using the PSAT 8/9. This limits us to 3rd and 4th grade students in the first year of each cohort. We also look at four-year growth through spring 2023 using 3rd graders in 2015 and 2019 and find smaller but still sizable effects.

[2] To ensure that we can observe all available benchmark and M-STEP exams for students between spring 2019 and spring 2023 and to align with the sample with the M-STEP growth analysis, we restrict to students in grades 5 and 6 in fall of 2020 This allows for a more consistent comparison of students and outcomes across assessments. Given that the available literature on learning during the COVID-19 pandemic generally finds that achievement slowed more for early elementary students than older students (e.g., Amplify Education, 2021; Pier et al., 2021; Goldhaber et al. 2022), this sample choice may imply our estimates show more achievement growth than we would see for lower grades.

math by fall 2020 than they were in reading. Achievement in both subjects then declined substantially between the fall and spring of 2020-21. Although students recovered some of these losses, further recovery appears to have stalled going into spring 2023, at a level well below national pre-pandemic norms. While we don't have pre-pandemic benchmark data in Michigan, which necessitates comparisons to national averages, these results are consistent with what we see in the less frequently assessed M-STEP data, where student scores remain below pre-pandemic achievement levels. Across both types of assessments, we consistently find larger negative estimates for students of color, students who are economically disadvantaged, and students whose districts did not offer in-person instruction in 2020-21.

The remainder of the paper proceeds as follows: Section two first describes Michigan's Return to Learn legislation that laid out assessment requirements to enable districts and policymakers to track student learning during the pandemic. Section three then briefly reviews the extant literature on student achievement during and beyond the pandemic. The fourth section describes our data and methods of estimating achievement growth and trends during the pandemic. We provide our results in the fifth section and conclude with a discussion of these results and implications for policymakers in section six.

2. K-12 Student Testing in Michigan during the COVID-19 Pandemic

In March of 2020, all schools in Michigan were ordered by the state to close and move to remote learning. The expected spring 2020 administration of the M-STEP exam was canceled, and schools stayed remote for the remainder of the school year. In August of 2020, the governor signed a series of three "Return to Learn" bills intended to grant districts flexibility to safely provide instruction during the COVID-19 pandemic (Public Act 147, 2020; Public Act 148,

2020; Public Act 149, 2020). For the 2020-21 school year only, the legislation waived many instructional requirements, including what learning activities count toward the attendance and enrollment calculations that determine state aid allocations. The state also waived requirements that students had to take M-STEP exams if they were in remote schooling. Approximately 70 percent of students participated in the M-STEP assessment in spring 2021, and the tested and untested populations differed substantially across individual, school, and district characteristics. As a result, given substantial sample selection concerns, we do not consider the spring 2021 administration of the M-STEP.

As a condition for receiving state aid for the year, the legislation required each district to develop an extended COVID-19 learning plan that included the administration of benchmark assessments to all K-8 students at the beginning and end of the school year to determine whether students made meaningful progress toward mastery of state standards in reading and mathematics. The Michigan Department of Education identified four benchmark assessments that are nationally normed, aligned with state content standards, and meet all other criteria outlined in state law.[3] The legislation allowed districts to choose an assessment from this state-approved list, another assessment meeting the same requirements, or develop their own assessment locally. While the legislation prohibited the use of these data for accountability purposes, districts that elected to use a state-approved provider were required to report data to the state. Additional legislation renewed the benchmark assessment requirement for the 2021-22 and 2022-23 academic years. In spring 2022, after nearly all schools in Michigan returned to full-

[3] The department required each approved assessment provider to prepare a transparency statement containing documentation verifying how their assessment(s) satisfy each of the criteria outlined in the law. These statements include information from alignment studies mapping benchmark assessment content to Michigan's content standards and linking studies demonstrating correlations between benchmark assessments and the M-STEP and mapping benchmark assessment scale score ranges to M-STEP proficiency levels.

time in-person instruction, the M-STEP exams returned to their pre-pandemic administration requirements and students were no longer given pandemic-related exemptions.

3. Relevant Literature

Across the country, educators and students alike have reported that teaching and learning during the pandemic were challenging, requiring educators to gain new skills, districts to provide new resources, and students to learn in unfamiliar and often difficult circumstances (e.g., Chen et al., 2021; Ferren, 2021; Francom et al., 2021; Hamilton et al., 2020; Pitluck & Jacques, 2021). Teachers, principals, and district superintendents reported that pandemic instruction was difficult for them and their students (Cummings et al., 2020; Hopkins et al., 2021). Survey evidence shows that educators were concerned that many students missed critical instructional time, had inadequate access to technology, lacked support for at-home learning, and received insufficient services during the 2020-21 school year (e.g., meals, counseling). In addition, educators indicated a need for training and guidance to help them provide adequate instruction during the pandemic. These challenges, combined with the extramural burdens of the pandemic, led to difficulties keeping students engaged in schoolwork, locating students, and maintaining student attendance (Cummings et al., 2020; Hopkins et al., 2021; for a review of the literature, see West & Lake, 2021).

It is therefore no surprise that a growing literature of national and state-specific research shows that there were fewer opportunities for students to learn during the pandemic than in a typical year. This has resulted in less – and sometimes far less – student growth on standardized achievement tests.

3.1. Student Achievement at the End of the 2022-23 School Year

As spring 2023 end-of-year assessment data have become available, there is growing evidence that students made progress academically during the 2021-22 school year, but such progress might have stalled in 2022-23 and many districts still fall below their pre-pandemic achievement levels, particularly in math. For example, in Tennessee, slightly more than a third of elementary, middle, and high school students scored proficient on the spring 2023 ELA standardized assessment. The scores for each grade span all matched or exceeded pre-pandemic ELA achievement levels. Math proficiency levels in Tennessee have yet to recover though as proficiency remains below 2019 levels (Tennessee Department of Education, 2022, 2023). As of spring 2022, state education agencies in Florida, Idaho, Indiana, Ohio, and Texas have all reported similar results (Appleton, 2022; Greater Fort Lauderdale Alliance, 2022; Kogan, 2022; Texas Education Agency, 2022; Idaho State Department of Education, 2022).

Analyses using nationally representative data from non-summative tests provide a more tepid view of pandemic recovery. A July 2022 study summarizing aggregate achievement among students who completed an NWEA assessment shows 2020-21 learning rates in math and reading were well below pre-pandemic trends. By 2021-22 achievement remained below pre-pandemic levels but there had been consistent improvement (Kuhfeld & Lewis, 2022). However, in 2022-23 there have been disturbing signs of stalled progress (Lewis & Kuhfeld, 2023).

Results from the spring 2022 administration of the National Assessment of Educational Progress (NAEP) paint an even bleaker picture of achievement during the pandemic. The most recent math and reading NAEP scores fell for nearly all student subgroups and in all regions across the country. On average, NAEP reading scores for students in grades four and eight dropped by three points relative to scores from 2019, which was the largest decrease in reading scores in more than 30 years. The declines in math were even larger (five and eight points for

4th- and 8th-graders, respectively) – the first time math scores fell since the NAEP began in the late 1960s (National Center for Education Statistics, 2022). Recently released results from the 2024 NAEP only show a tepid recovery since then in math and further degradation in reading (National Center for Education Statistics, 2025).

3.2. Heterogeneity in the Effects of the Pandemic on Student Learning

There are myriad reasons for these declines in student achievement, ranging from the massive toll the pandemic took on many educators' and students' mental, socio-emotional, and physical health, the frequent disruptions and changes to school operations, changes in learning environments and modes of instruction, and other extramural elements of the pandemic itself. A recent report from the Center on Reinventing Public Education (CRPE) detailed the overarching findings from the most rigorous studies of changes in these factors during the pandemic (Cohodes et al., 2022). The CRPE report highlights that many, and often the most traditionally underserved, students received less in-person instruction in the first two full school years affected by the pandemic than in a typical school year. This resulted in reduced learning time, and in some cases, lower quality instruction. This point is critical for any understanding of the effects of the pandemic on student learning. While average measures of interrupted learning are themselves quite concerning, it is clear from the CRPE's review that the effects of COVID-19 on students varied across student populations and the pandemic has had a greater, negative effect on achievement and achievement growth for specific student groups.

Relevant to this study, research consistently shows that Black, Latino, and economically disadvantaged students experienced the greatest learning interruptions and fell further behind their White and more advantaged peers (Amplify Education, 2021; Dorn et al., 2021; Goldhaber et al., 2023; Jack et al., 2022; Kilbride et al., 2022 Kogan & Lavertu, 2021; Pier et al., 2021). For

example, in the three metro-Atlanta districts studied by Sass and Ali (2022), differences in achievement by race and socioeconomic status have grown, more so in math than in reading.

Some of the variation in student achievement is also explained by the instructional modality districts used or students selected; students who received more in-person instruction, on average, learned more throughout the pandemic (Cohodes et al., 2022; Darling-Aduana et al., 2022; Jack et al., 2022; Kilbride et al., 2022; Kogan & Lavertu, 2021; Sass & Ali, 2022). For example, Goldhaber and colleagues (2023) leveraged NWEA assessment data from more than two million students across 49 states to understand how the provision of different instructional modalities impacted achievement gaps. Overall, math achievement gaps by race/ethnicity and school poverty status, as well as reading gaps to a lesser extent, did not widen in districts that provided students with in-person instruction. Conversely, the authors found that a district-level shift from in-person to remote instruction was a primary driver of widening racial/ethnic and socioeconomic achievement gaps.

With all these findings in mind, it is important to note that estimates of learning growth during the pandemic likely understate the true state of student learning. Across the country fewer students enrolled in school and absenteeism increased during the pandemic (Belsha, 2021; Cavitt, 2021; Levin, 2021; Mahnken, 2021; Pendharkar, 2021). This translates into lower-than-usual participation in assessments, especially in the 2020-21 school year, adding to the difficulty of drawing clear conclusions about student performance during the pandemic (Fensterwald, 2020; Sawchuk, 2021). Students disproportionately affected by the pandemic may comprise a substantial portion of the missing student assessment data, contributing to inequitable learning experiences across the country (Barnum, 2021).

4. Data and Methods

4.1 Data

We combine several sources of data to understand student achievement in Michigan during the COVID-19 pandemic, including student performance on both the state's summative end-of-year assessment and benchmark assessments administered during the pandemic. We also use state administrative data capturing student, school, district, and county demographics as well as a measure of access to in-person instruction offered during the 2020-21 school year. We describe these data below.

We use two sources of student achievement data to understand shifts in assessment performance during the pandemic. First, we use student outcomes from the M-STEP math and ELA assessments administered during the 2015-16 through 2022-23 school years. The M-STEP is Michigan's summative standardized assessment used to meet state and federal accountability requirements for students in grades three through seven. There are no M-STEP scores available from spring 2020, as the federal government waived testing requirements for the 2019-20 school year. Moreover, because the federal government waived test participation requirements in spring 2021 due to continued pandemic-related disruptions to in-person learning, only 73% of Michigan students participated and the tested population was not demographically representative of the student population. Hence, we drop the spring 2021 M-STEP from our data.

Second, we have student performance on nationally normed math and reading benchmark assessments administered to Michigan students in the fall and spring of the 2020-21 through 2022-23 school years at the district-grade-subgroup (e.g. economic disadvantage or race) level.[4]

[4] Although student-level data are not available for this analysis, researchers worked closely with the state and other stakeholders to develop a process to ensure that the aggregate data were constructed in a consistent manner across districts and reflect a stable sample of students across testing periods. Ninety-seven percent of districts allowed the

The vast majority of districts and students participated in either NWEA's MAP Growth or Curriculum Associates' i-Ready assessments. Due to the small sample sizes for the other two state-approved assessments, we limit our analyses to just districts that used MAP Growth or i-Ready. Our main outcome of interest for benchmark assessments is therefore district-grade-level average math and reading scores for students in grades 5 through 8 (grades 5 and 6 in 2020-21 through 7 and 8 in 2022-23), overall and by subgroups. A key benefit of this data is that it includes required testing during the first full "pandemic year" of 2020-21 when M-STEP was optional allowing us to track some of the early drops in achievement during the pandemic.[5]

Since we do not have pre-pandemic benchmark data, we use means and standard deviations from nationally representative norming samples to standardize scores for each grade, subject, and testing period. Therefore, we cannot use these data to identify pre-pandemic score distributions that are specific to our sample. Moreover, there are substantial differences between the MAP Growth and i-Ready samples in terms of demographic composition and prior achievement, and this approach also allows us to measure achievement on each benchmark assessment relative to comparable populations of students.

Although the M-STEP and benchmark data are not directly comparable, we include spring 2019, 2022, and 2023 M-STEP scores in our analysis of benchmark assessment trends to explore outcomes across assessments during a similar timeframe. While the M-STEP is not administered outside of Michigan, its design is closely based on the Smarter Balanced assessment and both M-STEP and Smarter Balanced scores are derived from the same

research team to access their student-level data and aggregate the data to the specifications needed for analysis, while the other 3% aggregated their data themselves using a template and set of instructions from the research team.

[5] Unfortunately, however, since these data start in fall 2020, we cannot observe what happens in the earliest phases of the pandemic. Nonetheless, as we will show there are important dynamics that continue throughout the first full pandemic year.

underlying scale (Michigan Department of Education, 2019). This allows us to convert M-STEP scores to Smarter Balanced scores and standardize outcomes relative to national norms for the Smarter Balanced assessment (Smarter Balanced Assessment Consortium, 2020) in analyses that use both assessments (figures 4 – 7 below).[6] Additionally, since the sample of students in our analysis were in grades 5 and 6 in the 2020-21 school year and 8th-graders in Michigan complete the PSAT 8/9 to satisfy annual federal testing requirements, we also standardize spring 2023 PSAT 8/9 scores for Michigan 8th graders relative to national norms.

Each testing regime has benefits and drawbacks, making it valuable to investigate both. For the M-STEP/PSAT, the data are recorded at the individual student level both before and after the start of the COVID-19 pandemic (for simplicity, we will refer to both exams as simply "M-STEP" going forward). This gives us the ability to control for the same characteristics included in the benchmark analysis at the individual student level rather than district-grade-subgroup averages and follow individual students throughout the period. Moreover, nearly all 3rd- through 8th-grade students in Michigan take the M-STEP, so these data provide a more representative and consistent measure of student achievement than the data from district-selected benchmark assessments, which only capture students who remained in the same district and received benchmark assessment scores during every possible testing period. However, since the M-STEP was not administered in spring 2020 and many students did not take the M-STEP in spring 2021, it is difficult to track student growth at different times throughout each pandemic-affected school year using that data.

[6] We first convert M-STEP scores to the theta scale following the scaling constants provided in Table 6-3 of Michigan's M-STEP Technical Report (Michigan Department of Education, 2019). Next, we convert to Smarter Balanced scores using the linear transformation shown in Section 5.2.3 of the Smarter Balanced Assessment Consortium's 2018-19 Summative Technical Report (Smarter Balanced Assessment Consortium, 2020). Finally, we standardize relative to the 2017-18 Smarter Balanced assessment national norms for comparability with the benchmark data.

The key benefits of the benchmark assessments begin with the fact that districts administer them twice each year, allowing us to examine higher frequency changes in achievement. This is particularly useful for seeing how far achievement dropped during the early pandemic before recovering. However, because these data are only available for fall 2020 and after, we cannot compare students' performance on these assessments directly to their pre-pandemic performance, nor can we fully capture changes in achievement during the earliest months of the pandemic between the spring and fall of 2020. Thus, while we use concordances between the M-STEP and one of the benchmark assessments to place them on the same scale relative to national norms, we caution that direct comparisons between the two types of exams are problematic given the different contexts under which they are administered and their differing purposes.

We merge assessment scores with several other data sources to explore heterogeneity in test score outcomes. First, we incorporate data on student demographic characteristics from the Michigan Student Data System (MSDS) to identify student subgroups based on their race/ethnicity and economically disadvantaged status.[7] In analyses exploring differences by race/ethnicity, we focus on White, Black, and Latino students as these are the three largest racial/ethnic subgroups in the state and we often do not have large enough sample sizes of students in other subgroups to permit analysis. Second, we examine heterogeneity by districts' instructional modality during the 2020-21 school year. In that year, all Michigan school districts not already operating virtually prior to the pandemic were required to report the instructional modalities offered to students each month of the school year. In the monthly questionnaire

[7] In Michigan, students are identified as economically disadvantaged if they qualify for free or reduced-price milk or meals through the National School Lunch Program (i.e., Supplemental Nutrition Eligibility). This includes homeless-identified students who are categorically eligible for free meals.

administered through MDE, districts were asked to indicate if they planned to instruct any of their students in a fully in-person (students receive 100% of their instruction in person), fully remote (students receive 100% of their instruction remotely), or hybrid format (students attend school in person for part of the week and participate in remote instruction for part of the week). For our analysis, we assign students to each modality type based on the number of months their district offered fully in-person instruction: zero months, one to four months, five to eight months, or all nine months of the 2020-21 school year.

Finally, since district modality offerings were often tied to community incidence of COVID-19, we link our achievement data with daily counts of county-level COVID-19 deaths collected and distributed by the Centers for Disease Control. This accounts for COVID-19 incidence in each district during our sample period. For our analysis of M-STEP outcomes, we average COVID-19 death rates throughout the 2020-21 school year and assign these rates to students in the pandemic cohort (COVID-19 death rates for students in the pre-pandemic cohort are set to zero). For the benchmark analysis, we assign death rates by averaging rates across the three months leading up to each test administration period (July, August, and September for the fall administration, and March, April, and May for the spring administration) in 2020-21,2021-22, and 2022-23.

4.2 Analytic Samples

M-STEP analysis. Our M-STEP analysis compares three-year M-STEP growth outcomes for two cohorts of students: pre-pandemic and pandemic. The pre-pandemic cohort includes approximately 198,600 students who completed the M-STEP math or ELA assessment in both spring 2016 and spring 2019. The pandemic cohort includes approximately 180,500 students who completed one iteration of the M-STEP math or ELA assessment prior to the pandemic in spring 2019 and again in spring 2022.

Given the three-year gap in outcomes and our desire to follow individual students, our analysis sample is constrained to include students who begin the three year-period in the 3rd- or 4th-grade and finish in the 6th- or 7th-grade.[8] Thus, the pre-pandemic cohort includes students who completed the 3rd- or 4th-grade assessment in 2016 and the 6th- or 7th-grade assessment in 2019. Similarly, students in the pandemic cohort include those who completed the 3rd- or 4th-grade assessment in 2019 and the 6th- or 7th-grade assessment in 2022.[9] Because we construct these measures only from students with data from both test administrations, we drop students who were not present in Michigan, did not participate in the test, or had invalid test scores in either period. Thus, the pre-pandemic and pandemic cohorts represent 91.7 and 83.4 percent of all Michigan 3rd- and 4th-grade students, respectively, who participated in M-STEP testing in the base year for each cohort.

Table 1 provides summary statistics for students in the M-STEP sample by subject and cohort. The table shows that students in the pre-pandemic and pandemic cohorts are similar demographically. More than half of the students in each cohort are female, and each cohort has similar shares of Black, Latino, special education, and English learner students. The two cohorts also started with similar base-year math and ELA achievement. However, we see that the pandemic cohort performed worse than the pre-pandemic cohort in math over the three-year period. Average math growth for students in the pre-pandemic cohort was essentially flat, while math achievement decreased by 0.212 SD on average for students in the pandemic cohort. Nonetheless, the difference in growth in ELA across cohorts was small; pandemic growth was only 0.03 SD smaller than pre-pandemic.

[8] In Michigan, 8th-graders take the PSAT 8/9 instead of the M-STEP, limiting us to examining students in grades three through seven when estimating growth models.

[9] We also conducted an analysis that looks at 4-year growth trends that includes 3rd graders in 2015 (7th graders in 2019) for the pre-pandemic cohort and 3rd graders in 2019 (7th graders in 2023) for the pandemic cohort. While the estimates for math are smaller (closer to zero) given the additional year of recovery, the demographic and remote learning patterns are similar. ELA results are similar across both timeframes. Results available upon request.

Benchmark Analysis. Our full sample for the benchmark analysis includes district-grade aggregated data from 97,733 students who entered the fall 2020 semester in grades five or six and have valid math or reading scores in all six administration periods between fall 2020 and spring 2023. As required by state law, these are aggregated to mean values at the district-vendor-grade-subject-semester level, both overall and within student subgroups based on race/ethnicity, gender, disability status, English learner status, and economic status. These aggregate measures only include students with test scores in all six semesters when benchmark assessments were administered (fall 2020 through spring 2023) to ensure that our comparisons over time reflect changes in student performance as opposed to changes in the populations of students tested. In total, this sample represents 68.4 percent of all 5th- and 6th-grade students in districts that offered a MAP Growth or i-Ready assessment in fall 2020.

Table 2 provides summary statistics for students in the benchmark assessment sample. In this table, we compare the characteristics of all 5th- and 6th-grade Michigan students ("Statewide" column) to those 5th- and 6th-grade students in the analytic sample who completed a MAP Growth or i-Ready assessment in fall 2020, spring 2021, fall 2021, spring 2022, fall 2022, and spring 2023 ("All," "MAP Growth," and "i-Ready"). While the demographics of students in the analytic sample generally resemble the statewide population – particularly in terms of achievement levels - they are less likely to be economically disadvantaged and/or Black.

4.3 Methods

To examine disparities in three-year M-STEP achievement growth between pre-pandemic and pandemic cohorts, we estimate the following baseline model:

$$3YG_{sgd} = \alpha + \theta_1 PCOHORT_{sgd} + \theta_2 BYA_{sgd} + \theta'_3 SCHAR_s + \gamma_g + \delta_d + \varepsilon_i \quad (1)$$

where $3YG_{sgd}$ represents three-year standardized M-STEP math or ELA growth for each student, *s,* in grade, *g,* and district, *d*. $PCOHORT_{sgd}$ is a binary indicator that identifies students in the pandemic cohort. BYA_{sgd} is the base year achievement for the student – 2016 in the pre-pandemic cohort and 2019 in the pandemic cohort. $SCHAR_s$ is a vector of student characteristics (i.e., gender and race/ethnicity, as well as economically disadvantaged, special education, English learner, unhoused, and migrant status). γ_g and δ_d are grade and district fixed effects. The coefficient θ_1 captures any disparity in standardized M-STEP test score growth between students in the pre-pandemic and pandemic cohorts. To estimate subgroup-specific differences in achievement growth during the pandemic, we extend model (1) by interacting $PCOHORT_{sgd}$ with our indicators for race/ethnicity, economically disadvantaged status, and access to in-person instruction (i.e., zero months, one to four months, five to eight months, and all nine months).

To understand trends in student achievement *during* the pandemic school years, we use both M-STEP and benchmark assessment scores in the following baseline model:

$$Y_{dgsvt} = \alpha + \theta_1 F20_t + \theta_2 S21_t + \theta_3 F21_t + \theta_4 S22_t + \theta_5 S22_t^{MSTEP} + \theta_6 F22_t + \theta_7 S23_t + \theta_8 S23_t^{MSTEP} + \theta'_9 DCHAR_{dgt} + \gamma_g + \delta_d + \varepsilon_i \quad (2)$$

where Y_{dgsvt} is the average standardized test math or reading score for students in district, *d*, grade, *g*, completing subject test, *s*, from assessment provider, *v*, in semester, *t*. *F20, S21, F21, S22, S22^{MSTEP}, F22, S23,* and *S23^{MSTEP}* are binary indicators identifying the semester associated with the outcome of interest, Y_{dgsvt} (i.e., nationally standardized M-STEP or benchmark assessment scores). $DCHAR_{dgt}$ is a vector of mean-centered, district-level student characteristics (i.e., student shares by gender and race/ethnicity, as well as economically disadvantaged, special education, English learner, homeless, and migrant status), and γ_g and δ_d are grade and district fixed effects, respectively. The coefficients

on indicators θ_1 through θ_8 describe the difference in average standardized test scores between spring 2019 (M-STEP) and fall 2020, spring 2021, fall 2021, spring 2022 (for both benchmark and M-STEP outcomes), fall 2022, and spring 2023 (for both benchmark and M-STEP outcomes) assessments, respectively. To examine heterogeneity across student subgroups and district instructional modality, we extend model (2) by interacting each time indicator with our indicators for race/ethnicity, economically disadvantaged status, and access to in-person instruction.[10]

5. Results

5.1. M-STEP Achievement Growth

We start by looking at trends in M-STEP achievement growth. The M-STEP allows us to look at achievement before and into the "late pandemic" period. However, unlike the benchmarks, it does not provide us with an understanding of the dynamics of achievement through the height of the pandemic. We use the benchmark assessments to look at that in the next section.

Figures 1 through 3 provide our results from estimating model (1), examining differences in 3-year achievement growth between students in the pre-pandemic and pandemic M-STEP cohorts. Tables A.1.1 through A.1.3 in the online appendix provide the coefficient estimates from these models. The zero-line represents the average three-year M-STEP growth for students in the pre-pandemic cohort overall or relative to the specific reference group. We show results from models that initially control for students' grade level then sequentially add demographic/community characteristics and district fixed effects.

Figure 1 shows that, overall, students in the pandemic cohort had significantly lower math achievement gains than students in the pre-pandemic cohort. Specifically, students in the pandemic cohort grew between 0.167 and 0.201 standard deviations less in math over the three pandemic-

[10] Since our measure of access to in-person instruction is calculated at the district level, we do not include district fixed effects in the models examining differences across modalities.

affected years than did students in the pre-pandemic cohort. ELA growth for students in the pandemic cohort was generally similar to growth in the pre-pandemic cohort; in our fully specified model, students who completed an ELA M-STEP assessment in 2019 and 2022 grew by approximately 0.025 standard deviations less than similar students who completed assessments in 2016 and 2019, however, this estimate is not statistically significant. Figure A.1.1 in the online appendix shows that the results are similar if we do not include district fixed effects or demographic and community control variables.

The second part of Figure 1 provides results examining heterogeneity by race/ethnicity and economically disadvantaged status. Even prior to the pandemic, disparities in achievement growth existed such that Black, Latino, and economically disadvantaged students experienced slower achievement growth than their White and higher-income peers. However, we find that growth disparities across these groups of students intensified during the pandemic, particularly in math. Specifically, in the three years prior to the pandemic, Black and Latino students experienced math achievement growth that was 0.116 and 0.018 sd lower than White students during the same period, respectively. In the three years encompassing the pandemic, Black and Latino achievement growth fell even further behind White students (-0.368 and -0.240 sd, respectively). Similarly, math achievement growth for economically disadvantaged students in the pre-pandemic cohort was 0.130 sd below their more advantaged peers and this disparity increased for students in the pandemic cohort (-0.351 sd). In ELA, achievement growth for Black, Latino, and economically disadvantaged students in the pre-pandemic cohort trailed their respective peers. However, these differences changed little over time for students in the pandemic cohort.

Figure 2 summarizes district fixed-effect models estimating differences in math and ELA M-STEP three-year growth by the instructional modalities provided to students in 2020-21.[11] We find that students in districts that offered in-person instruction all nine months of the 2020-21 school year still had lower math achievement growth over the course of the pandemic than students in the pre-pandemic cohort (-0.158 sd). Students in districts that did not offer in-person instruction for at least some of the 2020-21 school year experienced significantly slower math achievement growth than did students in districts that offered in-person instruction for all nine months, with achievement growth trailing their in-person peers by more than 0.05 sd. Moreover, achievement growth for these students trailed pre-pandemic students' math achievement growth by more than 0.2 sd. However, there were no significant differences between students in districts that were remote for all of the year or only part of it (i.e., in person for 5-8 months or for 1-4 months). Again, the disparities in ELA growth across modalities were much smaller, and the disparities in growth rates were not significant compared to students in the pre-pandemic cohort.

Finally, we find that learning remotely correlated with lower achievement for all students regardless of their race/ethnicity or socioeconomic status. Figure 3 shows results from models estimating differences in math and ELA growth by instructional modality provided to pandemic cohort students within each student demographic group considered in Figure 1. We find that the overall modality trends did not substantially differ across racial/ethnic and economically disadvantaged student subgroups, with all groups performing substantially higher in math and slightly higher in ELA if their school was in-person all year. For students experiencing remote instruction, Black and Latino students only showed slightly and mostly insignificantly lower math growth than White students with the same modality, as did economically disadvantaged students

[11] It is important to note that while we are considering three-year achievement growth covering 2019-20 through 2021-22 here, we only consider modality in 2020-21 as after the pandemic began in late 2020, all schools in the state were remote for the remainder of the school year and by fall 2021, almost every school district in the state had returned to in-person modality.

relative to non-disadvantaged.

5.2. Benchmark Achievement Trends

Figures 4 through 7 show adjusted trends in standardized math and reading benchmark achievement for students who started the 2020-21 school year in grades five or six and completed a MAP Growth or i-Ready assessment in all six administration periods during the 2020-21, 2021-22, and 2022-23 school years.[12] Since the trends in each specification are generally similar, we only report estimates for models that include district fixed effects.

As noted earlier, benchmark assessment scores are standardized relative to pre-pandemic national norms for each grade, subject, and testing period. As such, we interpret the trend lines in Figures 4 through 7 as deviations from the average scores for a nationally representative samples of students who took the same assessments before the pandemic. If Michigan students grew at the same rate as students in the pre-pandemic norming sample (and therefore maintained the same relative position within the norming distribution over time), we would see a straight horizontal line. If they grew at a faster rate than students in the norming sample, we would see lines that slope upward. By contrast, downward sloping lines indicate slower than expected growth between two time periods.

There are several important takeaways from Figure 4.[13] First, comparing the M-STEP estimates over the course of the pandemic, students fell substantially behind pre-pandemic levels initially with some recovery, but remain behind. By spring 2023 math and reading scores were still 0.07 and 0.08 SD below spring 2019, respectively.

While the M-STEP gives us some indication of where students ended up, the higher frequency of the benchmark assessments gives us the ability to track trajectories and assess how well

[12] Online Appendix Tables A.1.4 through A.1.6 provide the coefficient estimates from these models. Table A.1.4 summarizes overall math and reading benchmark trends and includes specifications that sequentially adds grade controls, district-level student controls and community-level COVID-19 incidence, and district fixed effects. Estimates without district fixed effects are similar and available by request.
[13] The coefficient estimates used to create the figure are provided in Online Appendix Table A.1.4.

the recovery is going. Achievement continued to drop in both math and reading benchmarks during the 2020-21 school year, falling even farther behind the national pre-pandemic norm. At the trough, math and reading scores on the benchmarks were around 0.1 standard deviations below national norms, and substantially below fall 2020 scores. Afterwards, students experienced recovery in math, with achievement slowly returning to the national norm. In reading, however, there was no recovery to speak of and the recovery in math appears to have stalled out as of fall 2022.[14] The M-STEP results in Spring 2022 and 2023 show the same stalled recovery in this later time period.

Thus, overall, achievement growth trends over the three years of the pandemic as measured by the benchmarks are consistent with our findings comparing pre- and post- pandemic M-STEP cohorts – a substantial drop in math achievement and a smaller drop in reading. What the benchmark assessments highlight, however, is that this path was non-linear with severe drops in the first fully impacted pandemic school year and some recovery in the time between spring 2021 and fall 2021 assessments. Math scores rose after that but appear to have stalled starting in fall 2022 with reading never recovering from the initial drop. What is particularly worrisome is that if one treats the benchmark and M-STEP exam scores as comparable, which we argue is reasonable given the similarities in the M-STEP and benchmark achievement levels in semesters when both are administered, recovery appears to have stalled at a level below pre-pandemic achievement. As such, the benchmarks show a clearer stalling pattern than the M-STEP given the higher frequency, providing a warning sign for continued recovery.

Figure 5 shows differences in adjusted trends in standardized math and reading benchmark achievement by race/ethnicity. Note that estimates here are standardized relative to the national

[14] To better see why a flat line indicates "normal" growth, Appendix Figures A.1.2 and A.1.3 show unadjusted scale score trends for the same sample of students. In these figures, the dashed gray lines represent pre-pandemic comparison points from each assessment provider's norming sample, and the solid blue and green lines represent math and reading outcomes for the cohorts of Michigan students tested during the pandemic. By comparing the slopes of the solid lines to the slopes of the dashed lines, we can see whether the score changes realized by Michigan students exceeded or trailed pre-pandemic norms. It is clear that in both math and reading the slopes between fall 2020 and spring 2021 of the solid lines are flatter than the dashed lines, indicating negative relative growth. This reverses in the next segment and then reverts in the last segment, though math remains parallel.

distribution using the Smarter Balanced 2018 national norms. We find similar patterns across subgroups, all in line with the overall results shown in Figure 4. White, Black, and Latino students all experienced a decrease in math and reading benchmark achievement between fall 2020 and spring 2021, followed by a rebound in scores during the 2021-22 school year that continued in the 2022-23 school year.[15] However, the initial drops from fall 2020 to spring 2021 are significantly larger for Black students in both subjects and for Latino students in math than for White students. Nonetheless, and surprisingly based on benchmark assessments, the recovery for Black students has been stronger than for White or Hispanic students, though this is perhaps because the initial drop was bigger. In the end, however, when looking at M-STEP scores, all three racial groups remain below pre-pandemic levels by roughly the same amount. As of Spring 2019 estimates in Online Appendix Table A.1.5 show that Black and Latino students scored about 0.5 SD and 0.25 SD lower than White students, respectively. Hence, even though the relative gap did not widen, the costs of the pandemic may have been more substantial for students in these groups as they had a considerably lower starting point.

Figure 6 examines similar trends across students who were and were not economically disadvantaged. We find many of the same trends as previously discussed. Economically disadvantaged students scored consistently lower in both math and reading across all testing periods compared to their more advantaged peers. Further, both groups of students experienced a decline in math and reading achievement between fall 2020 and spring 2021, but the drop was steeper for disadvantaged students falling by 0.08 SD and 0.05 SD more for disadvantaged students in math and reading, respectively.[16] As with the overall sample, achievement starts to recuperate starting after spring 2021 for both groups and exams but quickly stalls out in reading, never returning to baseline

[15] Estimates from Appendix Table A.1.5 show that overall, as expected, even when controlling for student and district characteristics, grade and district fixed effects, and COVID-19 death rates, Black and Hispanic achievement on both benchmarks and M-STEP are far below that for Whites. On average the estimates are 0.6 and 0.25 SD below White scores, respectively, over the sample period.

[16] See full set of coefficient estimates in online appendix table A.1.5.

while math scores continue to improve until fall 2022, at which point the recovery stalls for both disadvantaged and non-disadvantaged students.

Finally, Figure 7 shows differences in adjusted scale score trends in standardized math and reading benchmark achievement by 2020-21 instructional modalities. To clearly understand achievement trends among students in districts that offered varying amounts of in-person instruction, we have removed the confidence intervals from the benchmark exam markers in Figure 7 because they overlap to such a great extent, making the figure more difficult to interpret. Hence, it is important to note that the differences we see across modalities in Figure 7 are generally not statistically significant. Regression estimates with standard errors are provided in Online Appendix Table A.1.6.

First, we note that the extent to which a district offered in-person schooling in 2020-21 is clearly related to achievement levels. That is, ex-ante, schools that had lower achievement tended to remain in remote education longer. What is more important here, however, is how the temporal patterns differ. For both math and reading, the initial drop in achievement appears to increase with the number of months the district remained remote. In fact, in-person schools saw a slight increase in math scores between fall and spring 2021, though their spring 2022 M-STEP scores are below their spring 2019 scores, indicating that even schools that remained in-person throughout that year experienced learning delays. Nonetheless, these results show that in math, districts that maximized in-person schooling during the pandemic had the best performance and by 2023 their M-STEP math scores had nearly fully recovered. Reading, on the other hand, showed patterns that were more similar across modalities. While the initial drop remained larger for schools that had more time in remote instruction, the recovery was also sharper for these districts. Thus, consistent with the results in Figure 2, by the end of our period, modality does not appear to have a strong relationship with reading achievement growth through the pandemic.

These modality-related trends should be interpreted with caution, as decisions about instructional modality during the 2020-21 school year were related to pre-pandemic achievement levels. Districts that remained remote longer tended to have lower achievement levels even before the pandemic, and these decisions were likely partly shaped by longstanding systemic inequities. Prior research has shown that schools serving higher shares of economically disadvantaged students and minority students were more likely to remain in a remote learning environment due to higher levels of exposure to COVID-19, more limited resources, and other barriers (Parolin & Lee, 2021). As such, differences in achievement trends by modality may reflect not only variation in instructional access, but also deeper underlying inequalities.

We further explore trends in test scores by modality by separating the sample based on districts' cumulative COVID death rates per 100,000 people (as of May 2023) to see if the different responses to modality may instead be a function of overall COVID exposure. Online appendix figure A.1.5 shows that, while districts with above-median COVID death rates differ in their achievement levels, we see little evidence that the relationship between modality and achievement trends varied meaningfully by COVID exposure. However, our binary indicator for COVID exposure (above/below median) is only a proxy for the complex, multidimensional ways communities experienced the pandemic. It does not capture infection rates among students and staff, social and economic tolls, or other hardships caused by the pandemic.

6. Discussion

Our M-STEP results suggest that, while ELA achievement fell only slightly, math achievement growth dropped considerably during the pandemic relative to pre-pandemic cohorts. These decreases in achievement growth were larger for Latino and Black students than for White students, but there was no significant difference by race or ethnicity in ELA achievement growth over the same period. Similarly, economically disadvantaged students experienced larger

reductions in student achievement growth than their wealthier peers. In addition, students in districts that offered in-person instruction for all of the 2020-21 school year experienced significantly higher math achievement growth than those in districts that did not offer in-person instruction for part or all of the year.

Our benchmark results provide greater detail on student achievement trajectories during the two school years directly impacted by the pandemic and the initial stages of the recovery. In the first full pandemic-impacted school year (2020-21), achievement trends for Michigan 5th and 6th grade students fell further behind national pre-pandemic norms before partially rebounding during the 2021-22 school year, especially for math. However, the recovery has not been steep enough to return to pre-pandemic levels– as of spring 2023, both math and reading M-STEP scores remain below spring 2019 levels. Thus, going forward students would need to experience accelerated achievement growth – at rates greater than pre-pandemic expectations – to overcome the interrupted learning from the spring of 2020 and the 2020-21 school year. Unfortunately, benchmark assessments show stalled recovery since spring 2021 in reading and fall 2022 in math, indicating that for pandemic affected cohorts the learning losses may be permanent.

With these results in mind, we must advise caution when comparing results between the summative assessments and benchmark assessments. The exams are given under different environmental conditions and have different purposes – the former for school accountability, the latter explicitly for tracking student progress. Schools' and students' motivations to perform well on these assessments likely differ given the distinct purposes of the tests. Further, during 2021-22, benchmark exams were given regardless of remote status and thus were taken by students at home in districts that were remote, potentially inserting measurement error into the results during that year. One must also be cautious comparing these tests as they are all initially measured on

different scales. We attempt to mitigate this concern by using a pre-pandemic concordance of the M-STEP with the Smarter Balanced assessments and standardizing all scores relative to national norms on that exam. Importantly, while they are not strictly comparable to each other, scores on both types of exams in given years are relatively close to each other, suggesting that such comparisons are not unreasonable.

The overall patterns we see are consistent across all subgroups of students (by race/ethnicity and socioeconomic status). However, disparities in math achievement between White and Black or Latino students, as well between economically disadvantaged students and their wealthier peers, initially grew but then converged, such that they remain at a similar relative level, but at a lower base, at least in math. Finally, we find some evidence that students who had access to in-person instruction for the entirety of the 2020-21 school year performed better in both reading and math during that same school year, but these effects only persisted for math. Thus, larger deficits in math remain for districts that spent more time in remote instruction, despite starting at a slightly higher level in fall 2020.

7. Implications for Theory, Policy, and Practice

We make several recommendations for policymakers and educators based on these findings. First, results from the 2021-22 school year, and subsequent signs of the recovery stalling out make clear that the road to academic recovery will not be quick and a return to "business as normal" will be insufficient to return student achievement to pre-pandemic levels. Such a stalling out is consistent with national trends using less frequent assessment data. For instance, the National Assessment of Educational Progress (NAEP) shows large drops in

achievement in math and reading nationwide between 2020 and 2023 for 13-year-old students.[17] Our findings in Michigan should alert other states to the potential for stalling academic recoveries and suggest that it will be critical for states to watch for signs of speedbumps in pandemic recovery.

Moreover, our results and those from several other studies show particularly troublesome disruptions to math achievement. Yet, there has been relatively little discussion of ways to improve math achievement in the wake of the pandemic (Kuhfeld et al., 2022; Kuhfeld et al., 2022). While it is critical to continue providing supports for literacy instruction, the pandemic has taken an even greater toll on math achievement. Policymakers and educators will need to provide increased supports for math learning and instruction in the years to come.

In essence, our results from Michigan make clear that we are not "out of the woods" yet. Educators and policymakers must continue to monitor learning outcomes for all students, and especially for groups that were disproportionately affected by the COVID-19 pandemic. The mandated use and reporting of benchmark assessments in Michigan makes it possible for state and local policymakers to understand where progress is (and is not) being made towards academic recovery. It will be critical to continue collecting data that allow policymakers, educators, and stakeholders to assess progress in the coming years. In particular, research exploring trends in academic achievement over the past three years makes clear that the COVID-19 pandemic has had a greater and more negative effect on economically disadvantaged, Black, and Latino students. While we do find that outcomes for these students increased at a faster rate compared to their respective peers, disparities remain. These gaps were already large before the pandemic and hence, if anything, the pandemic may have temporarily halted progress on closing

[17] https://www.nationsreportcard.gov/ltt/?age=13

these disparities. Any decisions to reduce monitoring of student learning progress may exacerbate longstanding achievement gaps.

Beyond data collection and monitoring, we can draw from other literature to consider what policies and practices can help students recover from the pandemic learning disruptions we show here. Unfortunately, the most systematic work to date has shown limited success for more traditional interventions. Cabonari et al. (2024) study recovery impacts of tutoring and small group pull-out interventions, extra instruction time outside school hours, virtual learning programs, and extended school years in four large districts. While the expectation from earlier smaller scale studies is that these kinds of interventions could help, the number of students affected by COVID is so massive that it is infeasible to provide such intensive interventions at scale, leaving the scaled interventions to be ineffective. However, these interventions were done in 2021-22 when schools were still impacted by widespread disease and the immediate aftermath of remote instruction. It is possible that interventions like these could now be better targeted towards students who remain behind as others have recovered, improving implementation.

While effective policies to aid recovery for the COVID-19 pandemic cohort remain elusive, policymakers and practitioners may be better positioned to prevent negative impacts like those we find in Michigan in the event of another pandemic or large-scale localized disruptions to schooling due to natural disasters like hurricanes and major earthquakes. One implication of our findings and other studies of pandemic recovery is that students who spent more time in remote instruction had smaller learning gains. Our results show that this was particularly true for lower-income students. While these impacts need to be weighed against the potential costs to health in the case of a future pandemic or mass-scale learning disruption, policies that help keep students in-person should be prioritized. For example, masking appears to have been effective at

mitigating the health impacts of keeping kids in school during the COVID-19 pandemic (Cowger et al, 2022; Guzman et al, 2025). In a future respiratory disease pandemic caused by influenza or a different coronavirus, officials may consider masking in lieu of long-term school closures as an option provided medical research supports their effectiveness for the virus in question. For other types of disruptions, figuring out ways to return students to school quickly is imperative. Holding classes outdoors or using temporary spaces like vacant offices or community centers if school buildings are unusable for extended periods may be an effective mitigation practice.

In sum, our results bolster other data from around the country that make clear the road to recovery from COVID will be long – particularly for students who have been traditionally disadvantaged in K-12 public schooling. Educators and students will need continued and extensive supports to recover from the trauma of the COVID-19 pandemic, and governments at all levels must continue to prioritize both short- and longer-term investments into public education.

References

Amplify Education. (2021). COVID-19 means more students not learning to read. Amplify Education Research Brief. https://readytogether.sde.ok.gov/sites/default/files/2022-04/AmplifymCLASS_MOY-COVID-Learning-Loss-Research-Brief_022421.pdf

Appleton, A. (2022, July 13). Modest gains for this year's Indiana test scores after pandemic decline. Chalkbeat Indiana. https://in.chalkbeat.org/2022/7/13/23205866/ilearn-indiana-state-testingscores-2022-pandemic-recovery

Barnum, M. (2021, February 24). This year's state test results will be tough to make sense of, experts warn. Chalkbeat. https://www.chalkbeat.org/2021/2/24/22299804/schools-testing-covidresults-accuracy.

Belsha, K. (2021, August 7). Lack of in-person instruction pushed public school enrollment down, new research finds. Chalkbeat. https://www.chalkbeat.org/2021/8/7/22613546/research-remote-instructionschool-enrollment-declines.

Carbonari, M. V., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., ... & Staiger, D. O. (2024). The Impact and Implementation of Academic Interventions During COVID-19: Evidence from the Road to Recovery Project. AERA Open, 10, 23328584241281286.

Cavitt, M. (2021, July 12). Oakland County sees historic drop in public school enrollment during pandemic. The Oakland Press. https://www.theoaklandpress.com/2021/07/12/top-l-enrollment-0711.

Center for Educational Performance and Information, 2023. Student Enrollment Counts Report, Statewide Trend, All Grades K-12. https://www.mischooldata.org/student-enrollment-counts-report/.

Chen, L. K., Dorn, E., Sarakatsannis, J., & Wiesinger, A. (2021). Teacher survey: Learning loss is global—and significant. McKinsey & Co. https://www.mckinsey.com/industries/public-and-social-sector/our- insights/teacher-survey-learning-loss-is-global-and-significant?cid=other-eml-altmip-mck&hdpid=1a57cf47-eae2-400f-b9f0-fd8edddbbfb6&hctky=12238167&hlkid=b53175decdec48bbb6888bd21b5

Cohodes, S., Goldhaber, D., Hill, P., Ho, A., Kogan, V., Polikoff, M., ... & West, M. (2022). Student Achievement Gaps and the Pandemic: A New Review of Evidence from 2021-2022. Center on Reinventing Public Education.

Cowger, T. L., Murray, E. J., Clarke, J., Bassett, M. T., Ojikutu, B. O., Sánchez, S. M., ... & Hall, K. T. (2022). Lifting universal masking in schools—Covid-19 incidence among students and staff. New England Journal of Medicine, 387(21), 1935-1946.

Cummings, A., Kilbride, T., Turner, M., Zhu, Q., & Strunk, K. (2020). How did Michigan educators respond to the suspension of face-to-face instruction due to COVID-19. Education Policy Innovation Collaborative.

Curriculum Associates (2020, May). Scores on i-Ready Diagnostic that are equivalent to performance levels on the Michigan Student Test of Educational Progress (M-STEP). Curriculum Associates Research Report No. RR 2020-29.

Curriculum Associates. (2022, September). The state of student learning in 2022. Curriculum Associates Annual Report.

Darling-Aduana, J., Woodyard, H. T., Sass, T. R., & Barry, S. S. (2022). Learning-mode choice, student engagement, and achievement growth during the COVID-19 pandemic. AERA Open, 8.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2021, July 27). COVID-19 and education: The lingering effects of unfinished learning. McKinsey & Company.

Fensterwald, J. (2020, November 30). Early data on learning loss show big drop in math, but not reading skills. EdSource.

Ferren, M. (2021, July 6). Remote learning and school reopenings: What worked and what didn't. Center for American Progress.

Francom, G. M., Lee, S. J., & Pinkney, H. (2021, June 26). Technologies, challenges and needs of K-12 teachers in the transition to distance learning during the COVID-19 pandemic. TechTrends, 65(4), 589–601.

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.

Greater Fort Lauderdale Alliance. (2022, June 29). State releases 2022 Florida standards assessment and end of course exam results. The Alliance.

Guzman, M. A., Imberman, S. A., Filosa, N. R., Kilbride, T., & Malkus, N. (2025). Spillovers from Public Health Policies in Schools: Evidence from COVID Mask Mandates (No. w33849). National Bureau of Economic Research.

Halloran, C., Hug, C. E., Jack, R., & Oster, E. (2023). Post COVID-19 Test Score Recovery: Initial Evidence from State Testing Data (No. w31113). National Bureau of Economic Research.

Hamilton, L. S., Kaufman, J. H., & Diliberti, M. K. (2020). Teaching and leading through a pandemic: Key findings from the American Educator Panels spring 2020 COVID-19 surveys. Rand Corporation.

Hopkins, B., Turner, M., Lovitz, M., Kilbride, T., & Strunk, K. (2021) A look inside Michigan classrooms: Educators' perceptions of COVID-19 and K-12 school in the fall of 2020. Education Policy Innovation Collaborative.

Idaho State Department of Education. (2022, July 6). Early reading test shows statewide gains from previous spring and fall. [Press release].

Jack, R., Halloran, C., Okun, J., & Oster, E. (2023). Pandemic schooling mode and student test scores: evidence from US school districts. American Economic Review: Insights, 5(2), 173-190.

Kilbride, T., Hopkins, B., Strunk, K.O. & Yu, D. (2022) Michigan's 2020-21 and 2021-22 Benchmark Assessments. Education Policy Innovation Collaborative.

Kogan, V. (2022, June). Academic achievement and pandemic recovery in Ohio: An update from fall third grade English language arts assessments. Ohio State University: John Glenn College of Public Affairs.

Kogan, V., & Lavertu, S. (2021, January 27). The COVID-19 pandemic and student achievement on Ohio's third-grade English language arts assessment. Ohio State University: John Glenn College of Public Affairs, Ohio State University.

Kuhfeld, M., & Lewis, K. (2022). Student Achievement in 2021-2022: Cause for Hope and Continued Urgency. Collaborative for Student Growth. Brief. NWEA.

Kuhfeld, M., Soland, J., & Lewis, K. (2022). Test score patterns across three COVID-19-impacted school years. *Educational Researcher, 51(7), 500-506.*

Kuhfeld, M., Soland, J., Lewis, K., & Morton, E. (2022, March 3). *The pandemic has had devastating impacts on learning. What will it take to help students catch up?* Brookings.

Levin, K. (2021, March 18). Michigan lost 62,000 students this fall. Black enrollment fell 5%. Detroit Free Press.

Lewis, K., & Kuhfeld, M. (2023). Education's Long COVID: 2022-23 Achievement Data Reveal Stalled Progress toward Pandemic Recovery. Center for School and Student Progress at NWEA.

Mahnken, K. (2021, June 28). New federal data confirms pandemic's blow to K-12 enrollment, with drop of 1.5 million students; Pre-K experiences 22 percent decline.

Michigan Department of Education (2019). *Technical Report, Spring 2019 Michigan Student Test of Educational Progress (M-STEP).*

Michigan Public Act 147 of 2020, Mich., MCL § 388.1621f (2020)

Michigan Public Act 148 of 2020, Mich. MCL § 388.1701 (2020)

Michigan Public Act 149 of 2020, Mich., MCL § 388.1606 (2020)

National Center for Education Statistics (2022). NAEP long-term trend assessment results: Reading and mathematics. The Nation's Report Card.

National Center for Education Statistics (2025). NAEP long-term trend assessment results: Reading and mathematics. The Nation's Report Card.

NWEA (2020, December). Linking study report: Predicting performance on the Michigan state assessment system in grades 3-8 ELA and mathematics based on NWEA MAP Growth scores. NWEA Psychometric Solutions.

Parolin, Z., & Lee, E. (2021, April). Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature Human Behavior, 5(4), 522-528.*

Pendharkar, E. (2021, July 17). More than 1 million students didn't enroll during the pandemic. Will they come back? Education Week.

Pier, L., Hough, H. J., Christian, M., Bookman, N., Wilkenfeld, B., & Miller, R. (2021, January 25). COVID-19 and the educational equity crisis: Evidence on learning loss from the CORE Data Collaborative. PACE.

Pitluck, C., & Jacques, C. (2021, July). Persistent challenges and promising practices: District leader reflections on schooling during COVID-19. Research Brief. AIR. National Survey of Public Education's Response to COVID-19.

Sass, T., & Ali, S. M. (2022). Student Achievement Growth During the COVID-19 Pandemic: Spring 2022 Update.

Sawchuk, S. (2021, July 14). “Extreme” chronic absenteeism? Pandemic school attendance data is bleak, but incomplete. Education Week

Smarter Balanced Assessment Consortium (2020). *2018-19 summative technical report.*

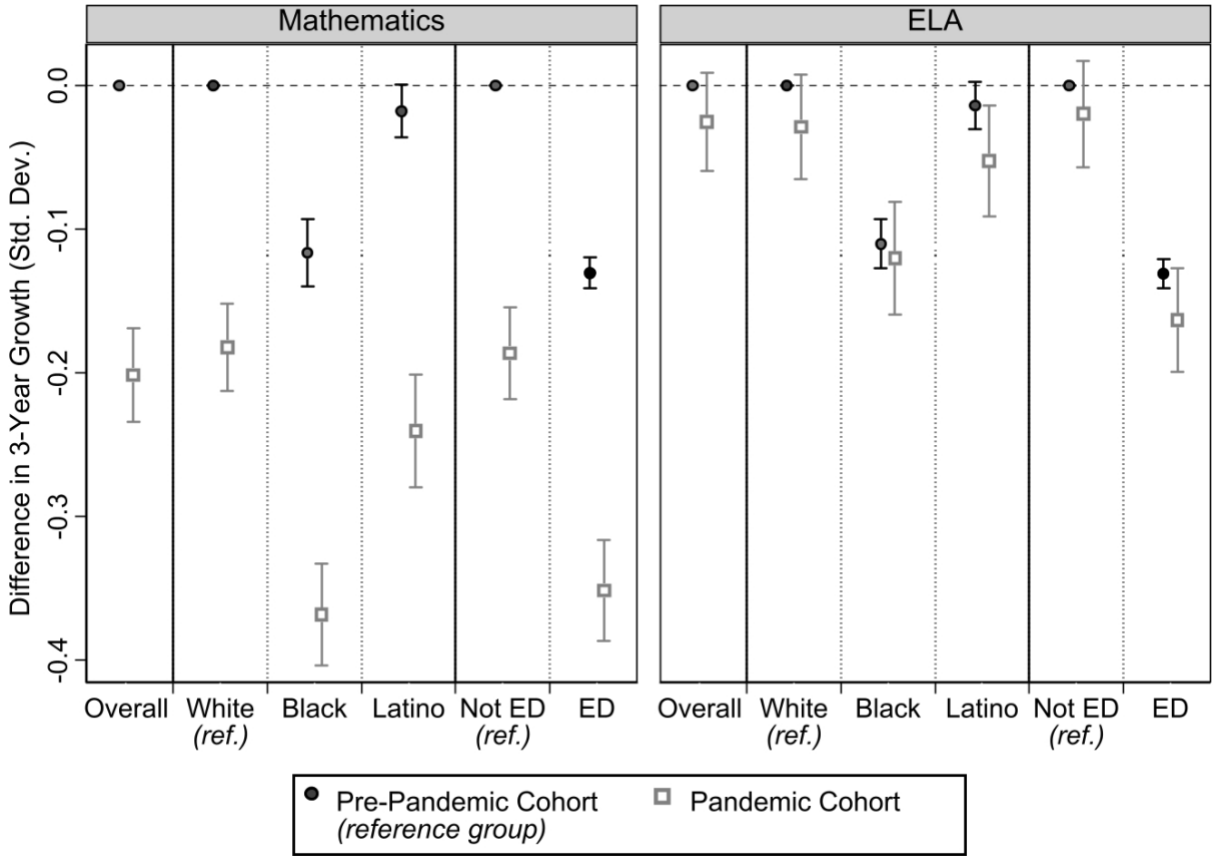
Tennessee Department of Education. (2022, June 14). Tennessee Releases 2021–22 TCAP State-Level Results Highlighting Significant Learning Acceleration.

Tennessee Department of Education. (2023, June 29). Tennessee Releases 2022–23 TCAP State-Level Results

Texas Education Agency. (2022, July 1). TEA Releases 2022 Grades 3–8 STAAR Results | Texas Education Agency.

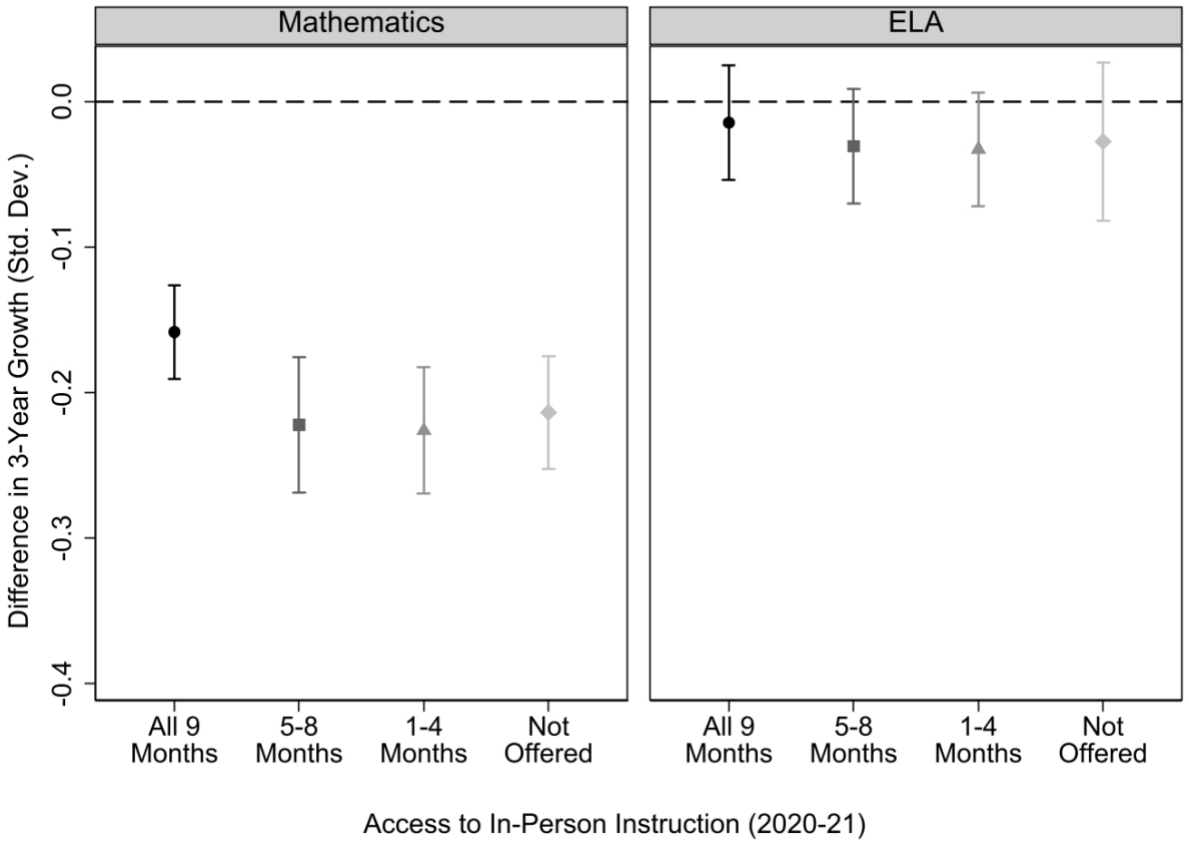
West, M. R. & Lake R. (2021, July). How much have students missed academically because of the pandemic? A review of evidence to date. Center on Reinventing Public Education.

Figure 1. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by Student Demographics, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



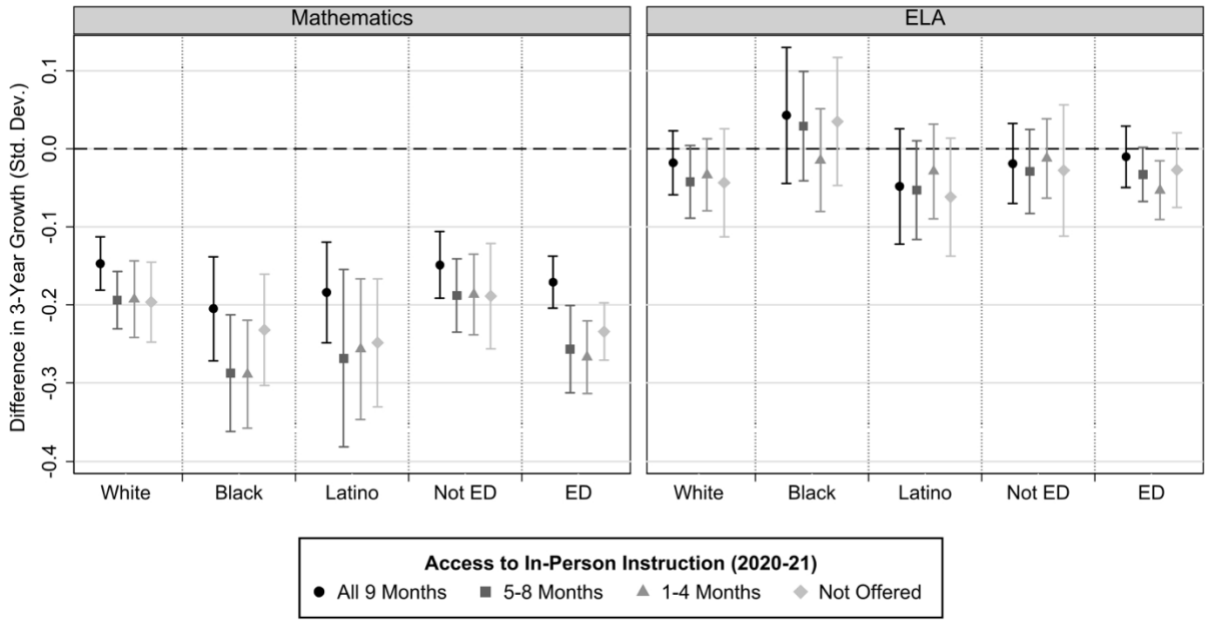
Notes: Each model includes student demographics and community characteristics, grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students, and district fixed effects to control for time-invariant, unobservable characteristics of each district that may influence learning trajectories.

Figure 2. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by 2020-21 Instructional Modality, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



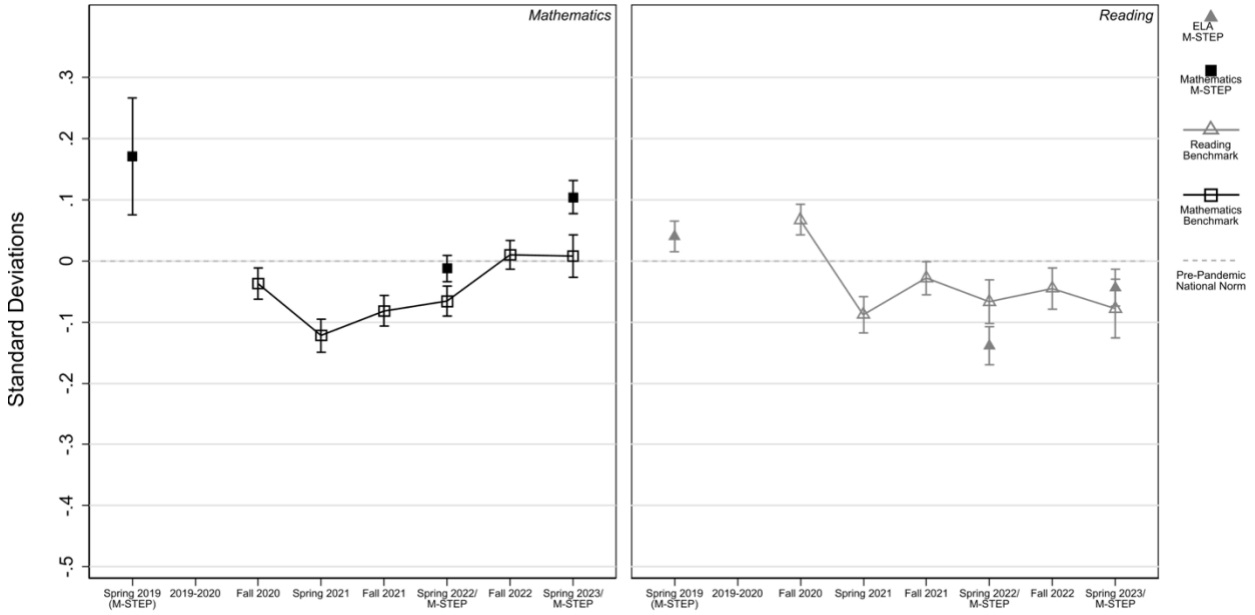
Notes: Each model includes student demographics and community characteristics, grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students, and district fixed effects to control for time-invariant, unobservable characteristics of each district that may influence learning trajectories.

Figure 3. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by 2020-21 Instructional Modality and Student Demographics, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



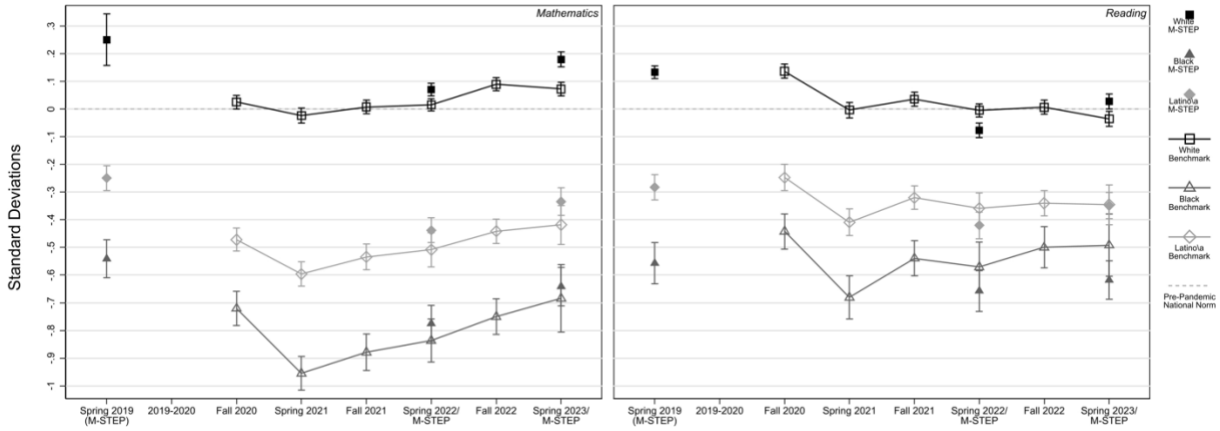
Notes: Each model includes student demographics and community characteristics, grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students, and district fixed effects to control for time-invariant, unobservable characteristics of each district that may influence learning trajectories.

Figure 4. Regression Adjusted Scale Score Trends, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6



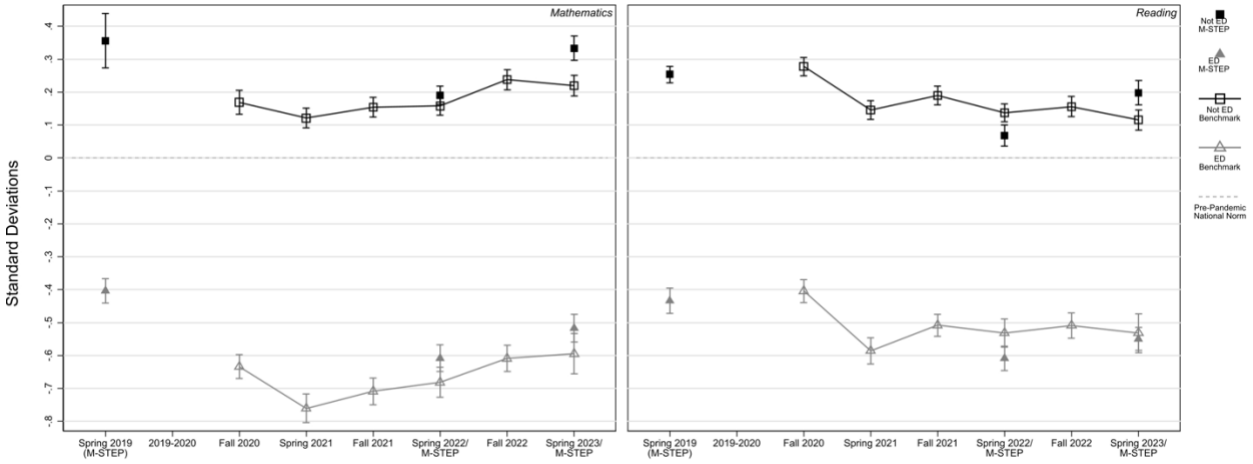
Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms.

Figure 5. Regression Adjusted Scale Score Trends by Race/Ethnicity, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6



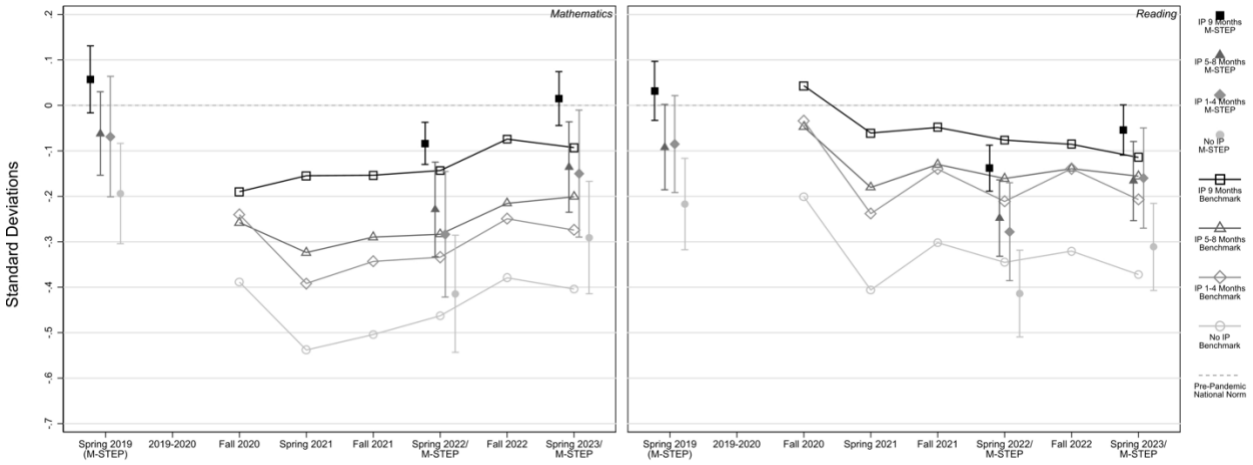
Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms.

Figure 6. Regression Adjusted Scale Score Trends by Economically Disadvantaged Status, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6



Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms.

Figure 7. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6



Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms.. Error bands have been removed from benchmark estimates for clarity of exposition.

Table 1. Summary Statistics; M-STEP Analytic Sample; Grades 3 and 4 (Base Year)

	Math Cohorts		ELA Cohorts	
	Pre-Pandemic	Pandemic	Pre-Pandemic	Pandemic
Total Students	198580	180573	198559	180499
Percent of all 3rd and 4th Grade in Base Year Included in Sample	91.7	83.4	91.7	83.4
Student Demographics (%)				
Economically Disadvantaged	52.1	52.8	52.1	52.8
Black	16.8	17.4	16.9	17.4
Latino	8.3	8.5	8.2	8.5
Special Education	11.3	12.3	11.3	12.3
English Learners	5.1	5.2	4.9	5.0
In-Person Access (%)				
9 Months	--	32.7	--	32.7
5-8 Months	--	31.8	--	31.8
1-4 Months	--	13.9	--	13.9
0 Months	--	21.6	--	21.6
M-STEP Scores (std. dev.)				
Base-Year Math Scores	0.0301	0.0364	--	--
Math Growth	0.0003	-0.2124	--	--
Base-Year ELA Scores	--	--	0.0268	0.0311
ELA Growth	--	--	-0.1061	-0.1376

Notes: Student demographic characteristics are measured in the comparison year for each cohort (i.e., 2019 for the pre-pandemic cohort and 2022 for the pandemic cohort. Base-year achievement summarizes outcomes in 2016 for the pre-pandemic cohort and 2019 for the pandemic cohort. "Math Growth" and "ELA Growth" represent three-year differences in achievement between 2016 and 2019 for the pre-pandemic cohort and between 2019 and 2022 for the pandemic cohort.

Table 2. Summary Statistics, Benchmark Assessment Analytic Sample, Grades 5-6 (2020-21)

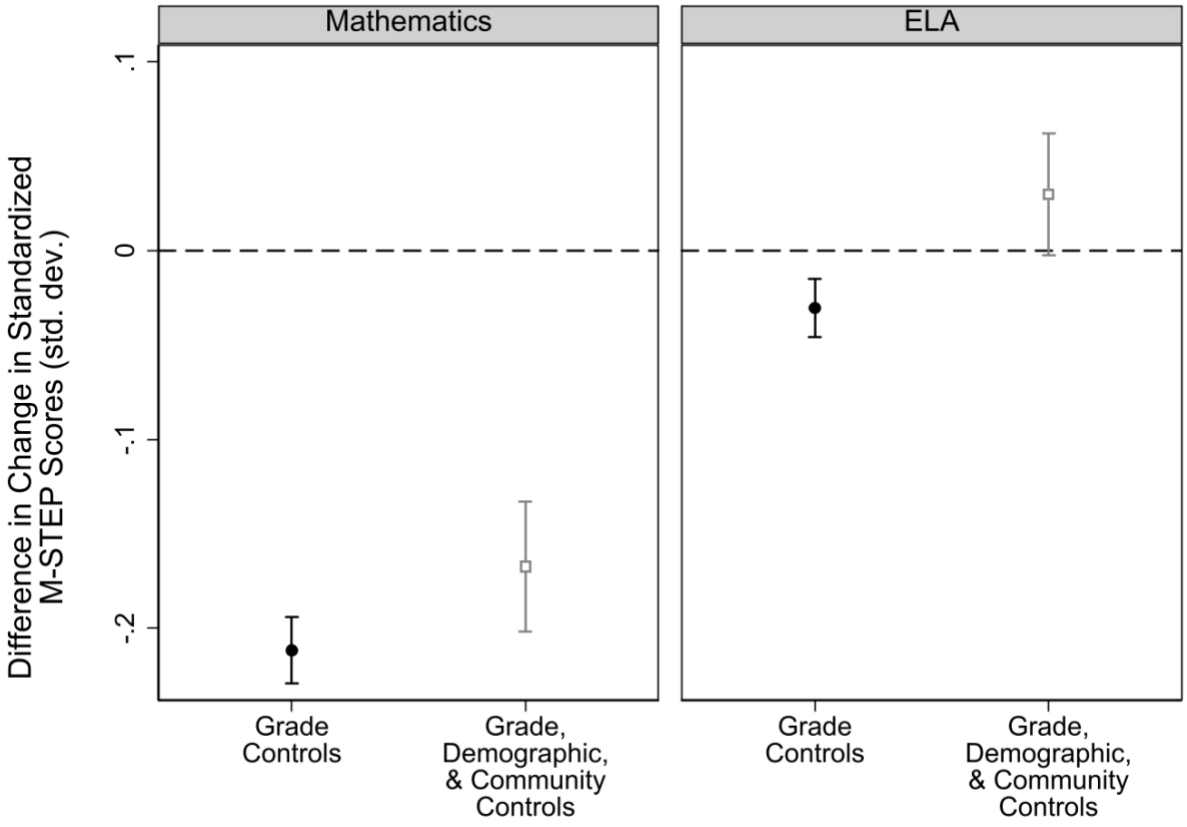
	Statewide	Analytic Sample		
		All	MAP Growth	i-Ready
Total Students	142929	97733	80434	17299
Percent of Analytic Sample	--	100.0	82.3	17.7
Percent of Enrollment in Districts Offering MAP Growth or i-Ready Assessment	100.0	68.4	56.3	12.1
Student Demographics (%)				
Economically Disadvantaged	53.9	48.7	47.6	53.6
Black	20.1	15.6	12.8	28.8
Latino	8.3	8.3	7.9	10.0
Special Education	12.9	12.5	12.5	12.5
English Learner	5.8	5.4	4.8	8.3
In-Person Access (%)				
9 Months	30.2	33.5	37.5	15.1
5-8 Months	32.1	32.2	26.1	60.5
1-4 Months	15.1	14.3	14.8	11.9
0 Months	22.6	19.9	21.5	12.5
2019 M-STEP Achievement (std. dev.)				
Math	-0.0627	0.0238	0.0554	-0.1247
ELA	-0.0896	-0.0100	0.0275	-0.1862

Notes: The "Statewide" column includes all 5th- and 6th-grade students in Michigan districts that offered an NWEA MAP Growth or Curriculum Associates i-Ready benchmark assessment. The "All" column includes both MAP Growth and i-Ready students from the analytic sample. Average standardized 2019 M-STEP achievement represents 3rd- through 5-th grade outcomes for all students in MAP Growth and i-Ready districts ("Statewide") as well as those in our analytic sample.

Online Appendix: Not for Publication

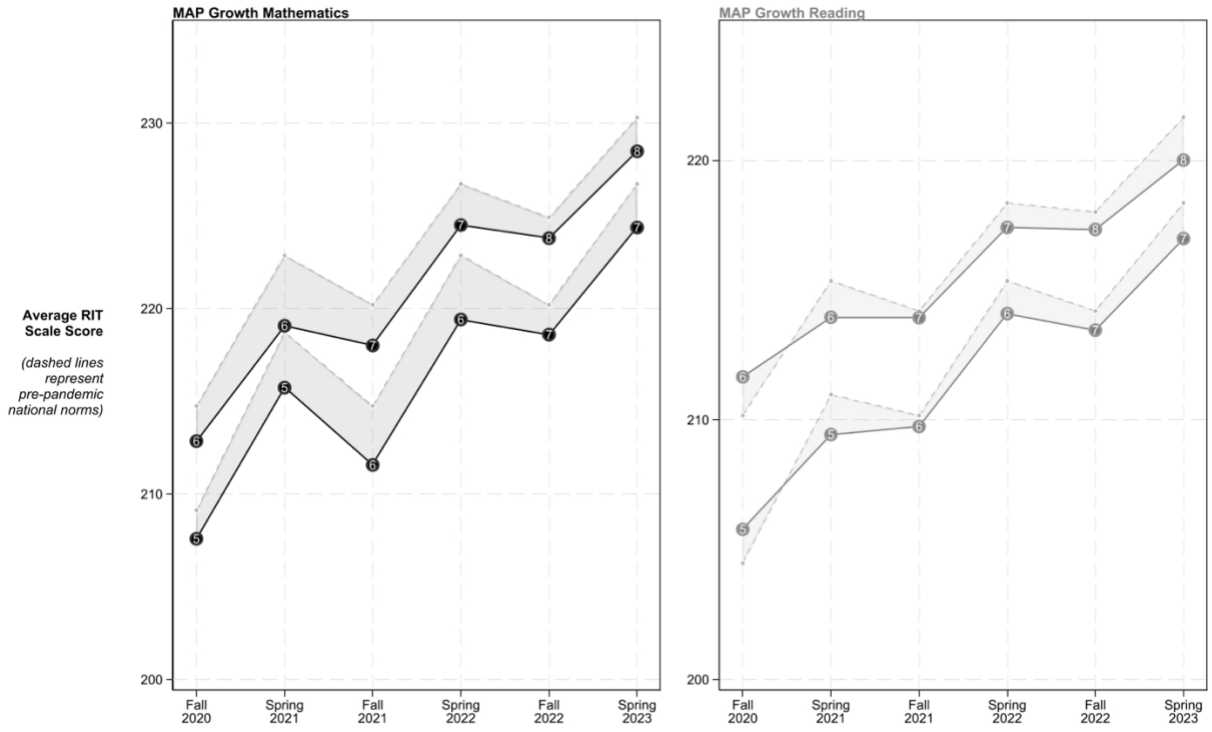
Appendix A.1. Supplementary Figures and Tables

Figure A.1.1. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



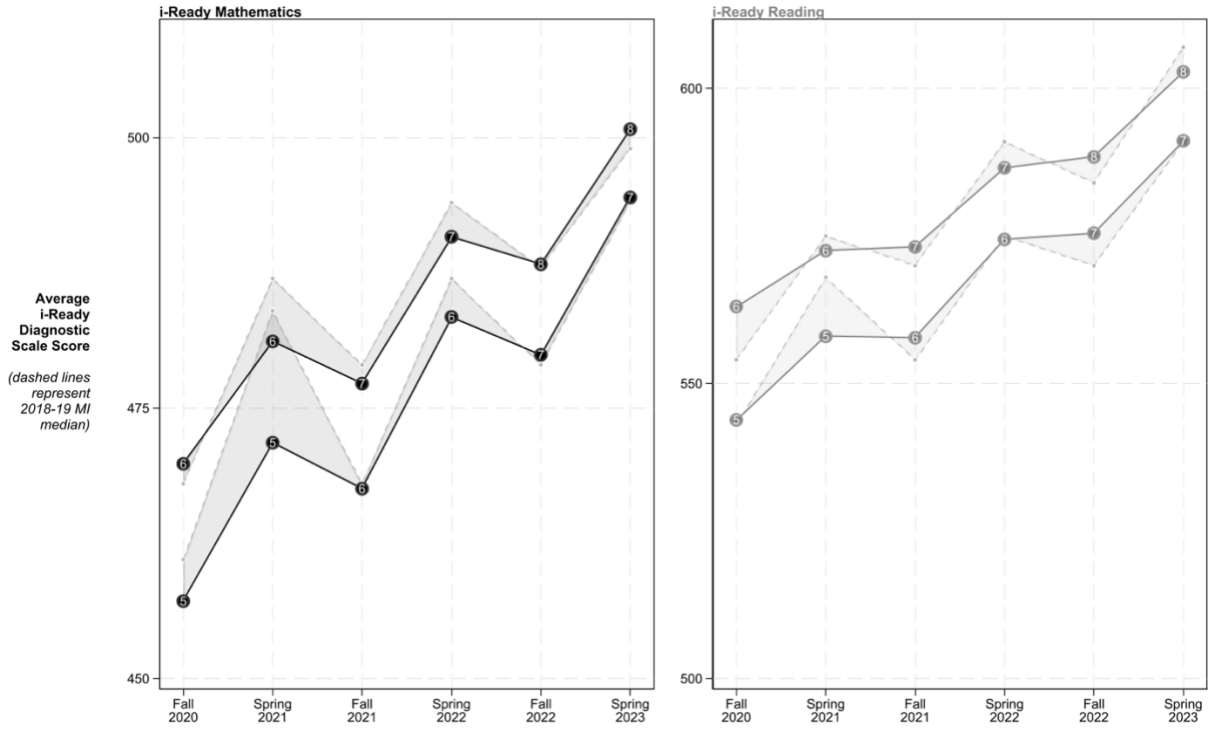
Notes: Each model includes grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students. The second estimate in each panel also includes controls for student demographics and community characteristics. The final estimate in each panel adds district fixed effects to control for time-invariant, unobservable characteristics of each district that may influence learning trajectories.

Figure A.1.2. Trends in Average Scale Scores, NWEA MAP Growth, Grades 5-7, Fall 2020 to Spring 2022



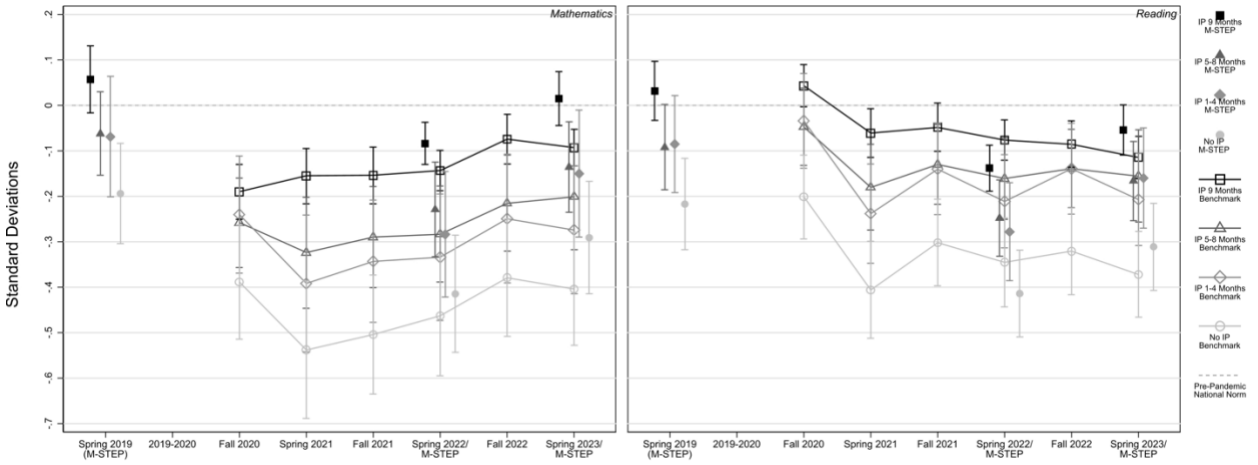
Notes: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent the 50th percentile of NWEA's conditional growth distribution. RIT stands for Rasch unit scale.

Figure A.1.3. Trends in Average Scale Scores, Curriculum Associates' i-Ready, Grades 5-7, Fall 2020 to Spring 2022



Notes: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent median scores for Michigan students in 2018-19.

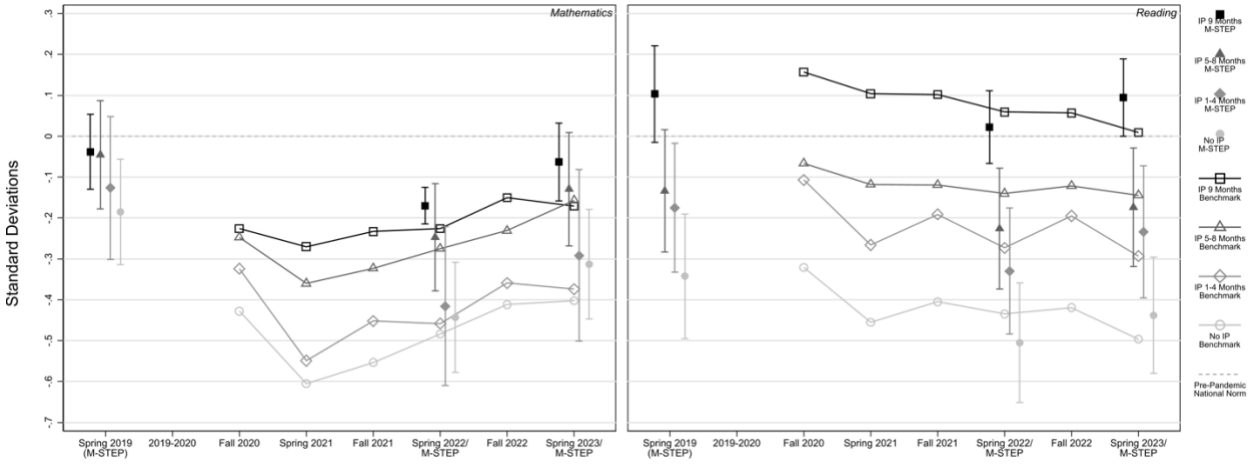
Figure A.1.4. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6



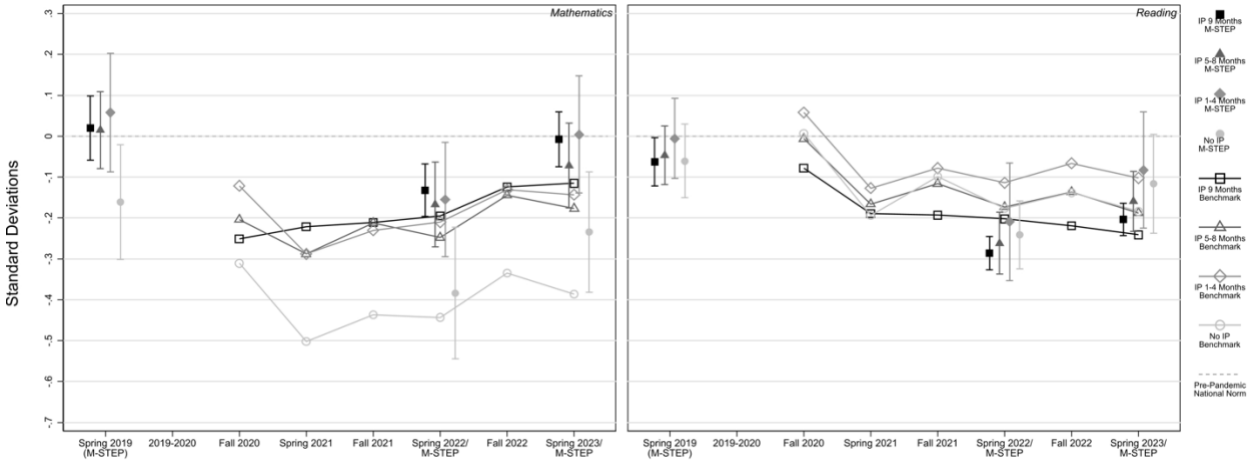
Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms..

Figure A.1.5. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality – By COVID Deaths Per 100,000 People, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6

Above Median COVID Deaths



Below Median COVID Deaths



Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to the 2018 Smarter Balanced assessment national norms. May 1st, 2023 used as cut-off for cumulative COVID deaths.

Table A.1.1. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments

	Mathematics			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort	-0.212*** (0.009)	-0.167*** (0.017)	-0.201*** (0.017)	-0.030*** (0.008)	0.030^{+} (0.016)	-0.025 (0.017)
Black		-0.188*** (0.010)	-0.150*** (0.007)		-0.071*** (0.010)	-0.101*** (0.007)
Latino		-0.069*** (0.010)	-0.037*** (0.007)		-0.034** (0.011)	-0.019* (0.008)
Economically Disadvantaged		-0.199*** (0.006)	-0.147*** (0.004)		-0.176*** (0.007)	-0.137*** (0.004)
Base-Year Achievement	-0.158*** (0.006)	-0.241*** (0.006)	-0.247*** (0.006)	-0.221*** (0.004)	-0.280*** (0.003)	-0.285*** (0.003)
Grade Controls	Y	Y	Y	Y	Y	Y
Student Controls	N	Y	Y	N	Y	Y
COVID-19 Death Rates	N	Y	Y	N	Y	Y
District Fixed Effects	N	N	Y	N	N	Y
R^2	0.091	0.145	0.190	0.105	0.140	0.177
N	379,153	379,153	379,153	379,058	379,058	379,058

Notes: Each model controls for student demographics and includes grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students. Robust standard errors clustered at the district level in parentheses. $^{+}$ *p < 0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table A.1.2. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by Student Demographics or 2020-21 Instructional Modality, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments

	Mathematics			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Cohort	-0.182*** (0.015)	-0.186*** (0.016)	-0.158*** (0.016)	-0.029 (0.019)	-0.020 (0.019)	-0.015 (0.020)
Black*Cohort	-0.070*** (0.017)			0.019 (0.012)		
Latino*Cohort	-0.041* (0.019)			-0.010 (0.013)		
Black	-0.116*** (0.012)			-0.110*** (0.009)		
Latino	-0.018^{+} (0.009)			-0.014^{+} (0.008)		
ED*Cohort		-0.035*** (0.010)			-0.013 (0.008)	
ED		-0.130*** (0.005)			-0.131*** (0.005)	
IP 5-8 Months*Cohort			-0.064** (0.020)			-0.016 (0.017)
IP 1-4 Months*Cohort			-0.068** (0.022)			-0.018 (0.020)
IP 0 Months*Cohort			-0.055*** (0.017)			-0.013 (0.024)
Base-Year Achievement	-0.247*** (0.006)	-0.247*** (0.006)	-0.246*** (0.006)	-0.285*** (0.003)	-0.285*** (0.003)	-0.285*** (0.003)
Student Controls	Y	Y	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y	Y	Y
COVID-19 Death Rates	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y
R^2	0.190	0.190	0.189	0.177	0.177	0.177
N	379,153	379,153	375,389	379,058	379,058	375,297

Notes: Each model controls for student demographics and includes grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students. Robust standard errors clustered at the district level in parentheses. $^{+}$ *p < 0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table A.1.3. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by 2020-21 Instructional Modality and Student Demographics, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments

	Math					ELA				
	White	Black	Latino	Non-ED	ED	White	Black	Latino	Non-ED	ED
Cohort	-0.147***	-0.205***	-0.184***	-0.149***	-0.171***	-0.018	0.043	-0.048	-0.019	-0.010
	(0.017)	(0.034)	(0.033)	(0.022)	(0.017)	(0.021)	(0.044)	(0.038)	(0.026)	(0.020)
IP 5-8 Months*Cohort	-0.047**	-0.083*	-0.084^{+}	-0.039*	-0.086***	-0.024	-0.014	-0.005	-0.010	-0.023
	(0.016)	(0.036)	(0.045)	(0.019)	(0.024)	(0.020)	(0.033)	(0.033)	(0.022)	(0.018)
IP 1-4 Months*Cohort	-0.046^{+}	-0.084*	-0.072^{+}	-0.038	-0.096***	-0.016	-0.058	0.019	0.007	-0.043*
	(0.026)	(0.036)	(0.044)	(0.027)	(0.023)	(0.023)	(0.035)	(0.035)	(0.024)	(0.021)
IP 0 Months*Cohort	-0.049*	-0.027	-0.064*	-0.040	-0.063***	-0.025	-0.008	-0.014	-0.009	-0.017
	(0.022)	(0.032)	(0.032)	(0.026)	(0.017)	(0.032)	(0.035)	(0.037)	(0.035)	(0.024)
Base-Year Achievement	-0.233***	-0.311***	-0.244***	-0.223***	-0.268***	-0.273***	-0.336***	-0.289***	-0.260***	-0.308***
	(0.003)	(0.017)	(0.006)	(0.003)	(0.008)	(0.003)	(0.005)	(0.005)	(0.003)	(0.003)
Student Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
COVID-19 Death Rates	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R^2	0.181	0.237	0.196	0.190	0.189	0.167	0.209	0.190	0.170	0.189
N	249,778	62,738	31,227	177,189	198,200	249,680	62,994	31,138	177,074	198,223

Notes: Each model controls for student demographics and includes grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students. Robust standard errors clustered at the district level in parentheses. $^{+}$ *p < 0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table A.1.4. Regression Adjusted Scale Score Trends, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6

	Mathematics			ELA		
	(1)	(2)	(3)	(4)	(5)	(6)
Fall 2020	-0.205***	-0.201***	-0.208***	0.027*	0.030*	0.027*
	(0.012)	(0.016)	(0.013)	(0.012)	(0.013)	(0.013)
Spring 2021	-0.284***	-0.269***	-0.293***	-0.128***	-0.118***	-0.128***
	(0.012)	(0.030)	(0.014)	(0.012)	(0.022)	(0.015)
Fall 2021	-0.247***	-0.239***	-0.252***	-0.068***	-0.063***	-0.068***
	(0.012)	(0.021)	(0.013)	(0.012)	(0.016)	(0.014)
Spring 2022	-0.232***	-0.224***	-0.237***	-0.107***	-0.102***	-0.107***
	(0.011)	(0.018)	(0.012)	(0.016)	(0.019)	(0.018)
Spring 2022 (M-STEP)	-0.178***	-0.171***	-0.183***	-0.179***	-0.174***	-0.179***
	(0.010)	(0.018)	(0.011)	(0.014)	(0.017)	(0.016)
Spring 2023	-0.161***	-0.158***	-0.163***	-0.118***	-0.116***	-0.118***
	(0.017)	(0.017)	(0.018)	(0.023)	(0.023)	(0.024)
Spring 2023 (M-STEP)	-0.064***	-0.061***	-0.066***	-0.084***	-0.082***	-0.084***
	(0.013)	(0.016)	(0.014)	(0.014)	(0.015)	(0.015)
i-Ready	-0.167	-0.022	-0.691**	-0.129	0.006	-0.085***
	(0.213)	(0.077)	(0.263)	(0.193)	(0.061)	(0.000)
Black, District Percent		-0.007	-0.243^{+}		0.038	-0.153
		(0.041)	(0.125)		(0.029)	(0.096)
Latino, District Percent		0.028	0.005		-0.001	0.033
		(0.037)	(0.040)		(0.026)	(0.028)
ED, District Percent		-0.204**	-0.028		-0.232***	-0.001
		(0.062)	(0.041)		(0.051)	(0.050)
Constant	0.074*	-0.035^{+}	0.348*	0.048^{+}	-0.048*	0.193^{+}
	(0.029)	(0.020)	(0.165)	(0.028)	(0.019)	(0.111)
Grade Controls	Y	Y	Y	Y	Y	Y
District-Level Student Controls	N	Y	Y	N	Y	Y
COVID-19 Death Rates	N	Y	Y	N	Y	Y
District Fixed Effects	N	N	Y	N	N	Y
R^2	0.052	0.523	0.908	0.036	0.533	0.883
N	8,055	8,037	8,037	7,956	7,938	7,938

Notes: Regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to national norms. $^{+}p < 0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table A.1.5. Regression Adjusted Scale Score Trends by Race/Ethnicity or Economically Disadvantaged Status, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6

	Math		Reading	
	(1)	(2)	(3)	(4)
Fall 2020	-0.225***	-0.187***	0.003	0.024+
	(0.012)	(0.018)	(0.013)	(0.014)
Spring 2021	-0.274***	-0.235***	-0.137***	-0.108***
	(0.014)	(0.015)	(0.014)	(0.014)
Fall 2021	-0.243***	-0.203***	-0.098***	-0.064***
	(0.012)	(0.015)	(0.013)	(0.015)
Spring 2022	-0.235***	-0.198***	-0.138***	-0.116***
	(0.011)	(0.014)	(0.012)	(0.014)
Spring 2022 (M-STEP)	-0.179***	-0.166***	-0.210***	-0.185***
	(0.012)	(0.014)	(0.013)	(0.016)
Spring 2023	-0.178***	-0.136***	-0.169***	-0.138***
	(0.013)	(0.016)	(0.014)	(0.016)
Spring 2023 (M-STEP)	-0.071***	-0.023	-0.106***	-0.055**
	(0.014)	(0.019)	(0.014)	(0.019)
Black	-0.540***		-0.556***	
	(0.035)		(0.038)	
Black*Fall 2020	-0.494***		-0.445***	
	(0.034)		(0.034)	
Black*Spring 2021	-0.680***		-0.542***	
	(0.031)		(0.040)	
Black*Fall 2021	-0.635***		-0.440***	
	(0.033)		(0.034)	
Black*Spring 2022	-0.601***		-0.432***	
	(0.037)		(0.045)	
Black*Spring 2022 (M-STEP)	-0.594***		-0.446***	
	(0.030)		(0.037)	
Black*Spring 2023	-0.504***		-0.322***	
	(0.060)		(0.057)	
Black*Spring 2023 (M-STEP)	-0.569***		-0.511***	
	(0.035)		(0.038)	
Latino	-0.250***		-0.283***	
	(0.023)		(0.023)	
Latino*Fall 2020	-0.248***		-0.251***	
	(0.020)		(0.023)	
Latino*Spring 2021	-0.323***		-0.272***	
	(0.021)		(0.022)	
Latino*Fall 2021	-0.292***		-0.223***	
	(0.021)		(0.019)	
Latino*Spring 2022	-0.274***		-0.221***	
	(0.028)		(0.025)	
Latino*Spring 2022 (M-STEP)	-0.259***		-0.210***	
	(0.019)		(0.021)	
Latino*Spring 2023	-0.241***		-0.177***	
	(0.033)		(0.035)	
Latino*Spring 2023 (M-STEP)	-0.263***		-0.243***	
	(0.022)		(0.023)	

	Math		Reading	
	(1)	(2)	(3)	(4)
ED*Fall 2020		-0.447*** (0.024)		-0.428*** (0.017)
ED*Spring 2021		-0.525*** (0.018)		-0.478*** (0.015)
ED*Fall 2021		-0.506*** (0.018)		-0.444*** (0.016)
ED*Spring 2022		-0.483*** (0.018)		-0.415*** (0.015)
ED*Spring 2022 (M-STEP)		-0.442*** (0.014)		-0.423*** (0.015)
ED*Spring 2023		-0.458*** (0.024)		-0.394*** (0.023)
ED*Spring 2023 (M-STEP)		-0.494*** (0.020)		-0.494*** (0.021)
i-Ready	-0.647* (0.257)	-0.602** (0.221)	-0.085*** (0.000)	-0.085*** (0.000)
Constant	0.333^{+} (0.187)	0.482** (0.163)	0.246* (0.124)	0.380*** (0.107)
District-Level Student Controls	Y	Y	Y	Y
Grade Controls	Y	Y	Y	Y
COVID-19 Death Rates	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
R^2	0.858	0.886	0.810	0.859
N	17,640	14,859	17,478	14,760

Notes: Regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for district-level student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to national norms. $^{+}$ *p < 0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table A.1.6. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-6

	Math	Reading
Fall 2020	-0.228*** (0.028)	0.008 (0.023)
Spring 2021	-0.216*** (0.021)	-0.104*** (0.021)
Fall 2021	-0.202*** (0.024)	-0.087*** (0.025)
Spring 2022	-0.189*** (0.019)	-0.115*** (0.021)
Spring 2022 (M-STEP)	-0.129*** (0.018)	-0.177*** (0.024)
Spring 2023	-0.128*** (0.020)	-0.148*** (0.023)
Spring 2023 (M-STEP)	-0.020 (0.029)	-0.088** (0.028)
IP 5-8 Months*Fall 2020	0.023 (0.031)	0.034 (0.030)
IP 5-8 Months*Spring 2021	-0.076** (0.027)	0.005 (0.027)
IP 5-8 Months*Fall 2021	-0.041 (0.032)	0.043 (0.031)
IP 5-8 Months*Spring 2022	-0.047^{+} (0.026)	0.039 (0.043)
IP 5-8 Months*Spring 2022 (M-STEP)	-0.053* (0.023)	0.015 (0.037)
IP 5-8 Months*Spring 2023	-0.018 (0.044)	0.081 (0.061)
IP 5-8 Months*Spring 2023 (M-STEP)	-0.060^{+} (0.033)	0.011 (0.035)
IP 1-4 Months*Fall 2020	0.048 (0.040)	0.040 (0.032)
IP 1-4 Months*Spring 2021	-0.135*** (0.030)	-0.059^{+} (0.031)
IP 1-4 Months*Fall 2021	-0.087** (0.033)	0.026 (0.034)
IP 1-4 Months*Spring 2022	-0.090* (0.036)	-0.016 (0.030)
IP 1-4 Months*Spring 2022 (M-STEP)	-0.099** (0.031)	-0.022 (0.041)
IP 1-4 Months*Spring 2023	-0.084* (0.034)	0.024 (0.034)
IP 1-4 Months*Spring 2023 (M-STEP)	-0.067 (0.046)	0.011 (0.049)
IP 0 Months*Fall 2020	0.024 (0.042)	0.004 (0.041)
IP 0 Months*Spring 2021	-0.160*** (0.035)	-0.096** (0.036)

	Math	Reading
IP 0 Months*Fall 2021	-0.124*** (0.035)	-0.004 (0.037)
IP 0 Months*Spring 2022	-0.095** (0.033)	-0.019 (0.034)
IP 0 Months*Spring 2022 (M-STEP)	-0.107*** (0.031)	-0.026 (0.038)
IP 0 Months*Spring 2023	-0.088* (0.036)	-0.009 (0.043)
IP 0 Months*Spring 2023 (M-STEP)	-0.083* (0.040)	-0.008 (0.047)
i-Ready	-0.691** (0.263)	-0.085*** (0.000)
Constant	0.352* (0.166)	0.195^{+} (0.112)
District-Level Student Controls	Y	Y
Grade Controls	Y	Y
COVID-19 Death Rates	Y	Y
District Fixed Effects	Y	Y
R^2	0.911	0.885
N	7,839	7,749

Notes: Regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for district-level student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019, 2022, and 2023 M-STEP estimates have been standardized relative to national norms. $^{+}$ *p < 0.10, * p<0.05, ** p<0.01, *** p<0.001*