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

May 2023

Education Policy Innovation Collaborative

COLLEGE OF EDUCATION | MICHIGAN STATE UNIVERSITY

236 ERICKSON HALL, 620 FARM LANE, EAST LANSING, MI 48824 | www.EPICedpolicy.org

DISCLAIMER

The Education Policy Innovation Collaborative (EPIC) at Michigan State University is an independent, non-partisan research center that operates as the strategic research partner to the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). EPIC conducts original research using a variety of methods that include advanced statistical modeling, representative surveys, interviews, and case study approaches.

This study used data the Michigan Education Research Institute-Michigan Education Data Center (MERI-MEDC) structured and maintained. MEDC data are modified for analysis purposes using rules governed by MEDC and are not identical to those data that MDE or CEPI collects and maintains. Results, information, and opinions solely represent the author(s) and are not endorsed by, nor reflect the views or positions of, grantors, MDE and CEPI, or any employee thereof. All errors are our own.

MAY 2023

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

AUTHORS

Katharine O. Strunk, *EPIC, Michigan State University*

Bryant Hopkins, *EPIC, Michigan State University*

Tara Kilbride, *EPIC, Michigan State University*

Scott Imberman, *EPIC Affiliated Faculty, Michigan State University*

Dongming Yu, *EPIC, Michigan State University*

ABSTRACT

Educators and policymakers have been concerned that the COVID-19 pandemic has 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 middle school students in Michigan. Comparing pre-pandemic to post-pandemic cohorts we find that math and ELA achievement growth dropped by 0.20, and 0.03 standard deviations more than expected, respectively, 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 followed by a partial recovery and potential stall-out in 2021-22.

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

INTRODUCTION

The COVID-19 pandemic has 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 continues 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). Similarly, National Assessment of Education Progress (NAEP) outcomes from spring 2022 represent historically large decreases in student achievement 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., 2022; Kilbride et al., 2022). In light of 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 the Michigan's summative end-of-year tests (the Michigan Student Test of Educational Progress, M-STEP) and formative 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 are able to examine both the total impact of the pandemic through spring 2022 as well as how achievement progressed *during* the pandemic-affected school years. 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). This paper answers three main questions:

1) How did the pandemic affect student achievement in Michigan?;
2) How did these achievement trends change throughout the pandemic?; and
3) Did achievement vary by race/ethnicity, economic disadvantage, and/or instructional modality?

To investigate M-STEP achievement growth, we compare three-year growth outcomes for a "pre-pandemic cohort" that completed either the math or ELA assessment three years apart before the school closures that occurred at the start of the pandemic (i.e., 3rd- and 4th-grade students in spring 2016 who progressed to 6th- and 7th-grade in spring 2019) and a "pandemic" cohort that completed the M-STEP before the pandemic and in the most recent test administration (i.e., 3rd- and 4th-grade students in spring 2019 who progressed to 6th- and 7th-grade in spring 2022). We also examine changes in achievement on nationally normed benchmark assessments across the fall 2020, spring 2021, fall 2021, and spring 2022 testing periods. 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 with the sample of students in our M-STEP analyses, we focus on middle school students (i.e., students who were in 5th through 7th grade in 2020-21 and in 6th through 8th grade in 2021-22). 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 also provide an upper bound for unfinished learning across all grade levels.

We find that middle school students in Michigan experienced far less math achievement growth over the last three years than prior cohorts of students before the pandemic. Effects on ELA were generally small and statistically insignificant. However, this overall picture of pandemic-era achievement masks semester-by-semester trends in achievement. In particular, Michigan students were scoring much farther behind national norms in math by fall 2020 than they were in reading. Both math and reading achievement then declined substantially between the fall and spring of 2020-21, with somewhat steeper declines in reading than in math. Although students have recovered some of these losses as of spring 2022, average scores in both subjects remain below pre-pandemic norms. Across both types of assessments, we consistently find larger negative effects 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.

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 legislation allowed districts to choose one of four state-approved benchmark assessments in reading or math, an assessment that met 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 academic year. Finally, in spring 2022, after nearly all schools in Michigan returned to full-time in-person instruction, the M-STEP exams returned to their pre-pandemic administration requirements and students were no longer given pandemic-related exemptions.

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). In Michigan, as well, 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 Michigan 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 things, 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.

Student Achievement at the End of the 2021-22 School Year

As spring 2022 end-of-year assessment data have become available, there is growing evidence that students made progress academically during the 2021-22 school year, but many still fall below 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 2022 ELA standardized assessment. The scores for each grade span all matched or exceeded pre-pandemic achievement levels. Math proficiency levels in Tennessee have yet to recover, though proficiency gains across all grade levels closed 30 to 50% of the initial learning gaps documented at the beginning of the COVID-19 pandemic (Tennessee Department of Education, 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 (Kuhfeld & Lewis, 2022). In 2021-22, learning gains generally mirror pre-pandemic achievement trends, and, in some cases, achievement growth exceeded that of a typical school year

by as much as a quarter to a third of the unfinished learning experienced throughout school closures and remote instruction over the last two school years. However, even if this accelerated growth continues at similar rates to those seen during the 2021-22 school year, it may be years before students experience a full recovery; Kuhfeld and Lewis (2022) estimate that students currently in grades three through five may not fully recover for three to five years while middle school students may need five or more years to return to pre-pandemic achievement levels.

Recently reported 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). Outcomes in some states were worse than others. In Michigan, where our study is based, NAEP math declines were generally equal to the average decreases across the country (four and eight points for 4th- and 8th-graders, respectively), but declines in reading scores for 4th- (six points) and 8th-graders (four points) exceeded national averages.

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, learning environments, 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 of these studies (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., 2022; 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, have 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 (2022) 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 of these findings in mind, it is important to note that estimates of learning growth during the pandemic likely understate the true impacts on student learning. Across the country and in Michigan, we know that fewer students enrolled in school and that 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). In particular, 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).

DATA AND METHODS

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, 2018-19, and 2021-22

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 in M-STEP testing in spring 2021, and the tested population consisted of more White, non-economically disadvantaged students from higher income districts with lower proportions of students of color. Given these limitations, our main M-STEP measures are generated as three-year changes in student M-STEP performance between 2016 and 2019 (for the pre-pandemic cohort) and changes between 2019 and 2022 (for the pandemic cohort). We use three-year gaps to ensure that we have a pre-pandemic testing outcome for the pandemic cohort. Prior to calculating these three-year growth outcomes, we standardize math and ELA M-STEP scores within each cohort to enable a comparison of student achievement over time. Specifically, we calculate the mean and standard deviation of math and ELA M-STEP scores separately for each grade level in the base year for each cohort (i.e., 2016 and 2019 for the pre-pandemic and pandemic cohorts, respectively). We then use these grade- and year-specific means to standardize math and ELA M-STEP scores for the same grade levels relative to the base year for each cohort.

Second, we use student performance on nationally normed math and reading benchmark assessments administered to Michigan students in the fall and spring of the 2020-21 and 2021-22 school years. 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 (Renaissance Learning's Star 360 and Data Recognition Corp's Smarter Balanced Interim Assessments), we limit our analyses to just MAP Growth and i-Ready. Due to Michigan policies written into the Return to Learn law, we are restricted to using district-grade-subgroup level means rather than individual student data.[1] Our main outcome of interest for benchmark assessments is therefore district-level average math and reading scores for students in grades five through seven, overall and by subgroups.

Similar to the M-STEP outcomes, the benchmark assessment scores are standardized relative to pre-pandemic test score distributions. However, unlike the M-STEP outcomes, we use means and standard deviations from nationally representative norming samples to standardize scores for each grade, subject, and testing period. One reason for this is that districts only provided benchmark assessment data from the fall 2020 and subsequent testing periods because they were not required to administer assessments prior to fall 2020, and for those that did, they were not required to provide them to the state. 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 populations of students that are more comparable to each other.

Although the M-STEP and benchmark data are not directly comparable, we include spring 2019 and 2022 M-STEP scores in our analysis of benchmark assessment trends to explore outcomes across assessments across a similar timeframe. While the M-STEP is not administered outside of Michigan, its design is based closely on the Smarter Balanced assessment and both M-STEP and Smarter Balanced scores are derived from the same 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). Additionally, since the sample of students in our analysis were in grades 5 through 7 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 2022 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, the data are recorded at the individual student level both before and after the start of the COVID-19 pandemic. 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. Moreover, nearly all 3rd- through 7th-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. 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. As such, we use three-year periods to measure achievement growth and the pandemic cohort includes some instruction in 2019 before the start of the pandemic.

The key benefits of the benchmark exams begin with the fact that districts administered them twice each year, allowing us to examine higher frequency changes in achievement. For example, with the benchmark exams we can study how far achievement initially fell over the course of the first full pandemic year, and then how quickly students recovered. 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 capture changes in achievement during the earliest months of the pandemic between the spring and fall of 2020. In addition, there is a national sample of students who take the NWEA MAP Growth and Curriculum Associates i-Ready assessments. This more easily allows us to compare Michigan students' progress throughout the pandemic to that of students across the country.

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.[2] 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 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 Michigan Department of Health and Human Services in order to control for COVID-19 incidence during our sample period. We use these data to calculate seven-day average death rates per 100,000 residents for the first day of each month between July 2020 and May 2022. 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 equal 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 both 2020-21 and 2021-22.

Analytic Samples

M-STEP analysis

Our M-STEP analysis compares three-year M-STEP growth outcomes for two groups of students: our pre-pandemic and pandemic cohorts. The pre-pandemic math and ELA cohorts include approximately 198,600 students who completed the M-STEP math or ELA assessment in both spring 2016 and spring 2019. The pandemic math and ELA cohorts include approximately 180,500 students who completed one iteration of the M-STEP math or ELA assessment prior to the pandemic in spring 2019 and the most recent administration of the assessment in spring 2022. The difference in size between

the pandemic and pre-pandemic cohorts is likely due to the fact that K-12 student enrollment in Michigan has decreased each year over the last decade, with particularly acute declines in 2020-21 (Center for Educational Performance and Information, 2023).

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.[3] 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. 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 88.0 and 85.3 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. Table 1 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, previewing our results, we see that the pandemic cohort performed worse than the pre-pandemic cohort over the three-year period. Average math growth for students in the pre-pandemic cohort was small but positive (0.030 standard deviations [sd]), while the math achievement decreased by 0.212 sd on average for students in the pandemic cohort. Students in the pre-pandemic and pandemic cohorts experienced a similar decrease in ELA achievement over their respective three-year periods.

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 3rd and 4th Graders in Base Year	88.0	85.3	88.0	85.3
STUDENT DEMOGRAPHICS (%)				
Economically Disadvantaged	52.1	52.7	52.1	52.8
Black	16.8	17.4	17.0	17.4
Latino	8.3	8.5	8.2	8.5
Special Education	11.3	12.3	11.3	12.3
English Learners	8.2	9.2	8.0	9.1
IN-PERSON ACCESS (%)				
9 Months	--	32.6	--	32.6
5-8 Months	--	31.9	--	32.0
1-4 Months	--	13.8	--	13.8
0 Months	--	21.6	--	21.5
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.

Benchmark Analysis

Our full sample for the benchmark analysis includes district-grade aggregated data from 141,034 students who entered the fall 2020 semester in grades five through seven and have valid math or reading scores in all four administration periods between fall 2020 and spring 2022. We focus only on students in grades five through seven to provide the closest comparison to students in our M-STEP sample. These aggregate measures only include students with test scores in each of the four semesters when benchmark assessments were administered (fall 2020, spring 2021, fall 2021, and spring 2022) to ensure that our comparisons over time reflect changes in student performance as opposed to changes in the populations of students tested. Additionally, we exclude districts that were not required to report data under Michigan's benchmark assessment legislation. In total, this sample represents 68.8 percent of all 5th- through 7th-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- through 7th-grade Michigan students ("Statewide" column) to those 5th- through 7th-grade students in the analytic sample who completed a MAP Growth or i-Ready assessment in fall 2020, spring 2021, fall 2021, and spring 2022 ("All," "MAP Growth," and "i-Ready"). While the demographics of students in the analytic sample generally resemble the full population of students in districts that offered a MAP Growth or i-Ready assessment in similar grade levels, they are considerably less likely to be economically disadvantaged and slightly less likely to be Black. Students who took the NWEA MAP Growth assessment represent more than 80 percent of the analytic sample and have about 10 percentage points lower economic disadvantage rates and 6.5 percentage points fewer Black students than the average district statewide. Students who took the i-Ready assessments, on the other hand, are substantially more likely to be Black (30 percent) compared to the full population of Michigan students in grades five through seven (20 percent), but largely similar otherwise. This is largely driven by the Detroit Public Schools Community District, which is the largest school district in Michigan and accounts for more than one-fifth of all students who took an i-Ready assessment.

Table 2. Summary Statistics, Benchmark Assessment Analytic Sample, Grades 5-7 (2020-21)				
	Statewide	Analytic Sample		
		All	MAP Growth	i-Ready
Total Students	205,038	141,034	116,015	25,019
Percent of Analytic Sample	--	100.0	82.2	17.7
Percent of Enrollment in Offering MAP Growth or i-Ready Assessment	100	68.8	56.6	12.2
STUDENT DEMOGRAPHICS (%)				
Economically Disadvantaged	54.4	46.6	45.0	54.2
Black	20.0	16.6	13.5	30.2
Latino	8.8	8.3	8.0	9.7
Special Education	14.4	12.1	12.0	12.3
English Learner	5.3	4.8	4.1	8.5
In-Person Access (%)				
9 Months	31.9	38.3	38.8	32.8
5-8 Months	33.7	27.5	27.1	31.9
1-4 Months	14.9	15.3	16.1	6.0
0 Months	19.4	19.0	18.1	29.3
2019 M-STEP ACHIEVEMENT (STD. DEV.)				
Math	0.0011	0.0413	0.0745	-0.1121
ELA	0.0131	0.0268	0.0599	-0.1262

Notes: The "Statewide" column includes all 5th- through 7th-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 5th-grade outcomes for all students in MAP Growth and i-Ready districts ("Statewide") as well as those in our analytic sample.

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. $SCHAR_s$ is a vector of student characteristics (i.e., gender and race/ethnicity, as well as economically disadvantaged, special education, English learner, homeless, 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).

Second, 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:

$$Y_{dgsvt} = \alpha + \theta_1 S19_t^{MSTEP} + \theta_2 S21_t + \theta_3 F21_t + \theta_4 S22_t + \theta_5 S22_t^{MSTEP} + \theta_5' DCHAR_{dgt} + \gamma_g + \delta_d + \varepsilon_i \quad (2)$$

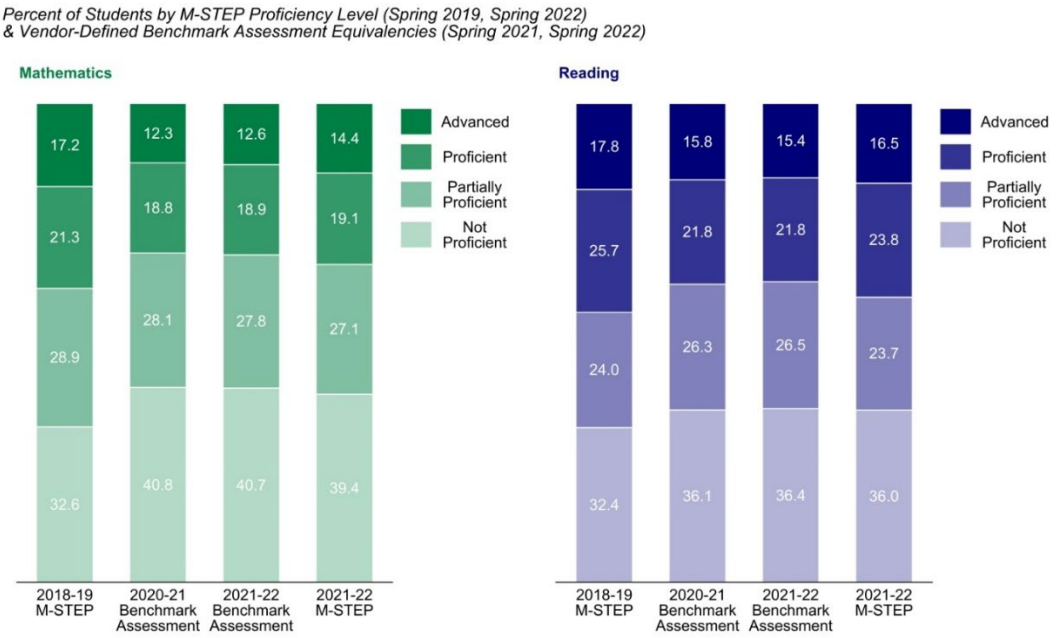
where Y_{dgsvt} is the average standardized test math or reading score for students in district, *d*, grade, *g*, completing subject test, *s*, from assessment provider, *v*, in semester, *t*. $S19^{MSTEP}, S21, F21, S22$ and $S22^{MSTEP}$ are binary indicators identifying the semester associated with 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, θ_2, θ_3, θ_4 and θ_5 describe the difference in average standardized test scores between fall 2020 and spring 2019, spring 2021, fall 2021, and spring 2022 (for both benchmark and M-STEP outcomes), 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.[4]

RESULTS

Before delving into the main sets of results, we first take a simple descriptive look at Michigan student performance over the course of the pandemic. Using linking studies available from each assessment provider (see Curriculum Associates, 2020; NWEA, 2020), we translate students' MAP Growth and i-Ready scores into approximate M-STEP proficiency levels (i.e., not proficient, partially proficient, proficient, or advanced) to investigate how Michigan students' benchmark assessment scores translate to M-STEP performance before and during the COVID-19 pandemic. For this analysis, we compare these performance thresholds for all 3rd- through 7th-grade students with a valid MAP

Growth or i-Ready benchmark assessment score in spring 2021 or spring 2022 to the actual distribution of 2018-19 and 2021-22 M-STEP proficiency outcomes among students in the same districts that offered a MAP Growth or i-Ready assessment. This analysis allows us to understand how Michigan students might have performed on the state's summative assessment during the first two pandemic years when M-STEP was either canceled or optional and compare these estimates to the actual M-STEP proficiency levels of students in the same districts in 2018-19 and 2021-22.

Figure 1. M-STEP Proficiency Levels and Vendor-Defined M-STEP Equivalencies, NWEA MAP Growth and Curriculum Associates i-Ready



Notes: These percentages include 3rd- through 7th-grade students with a valid benchmark assessment score in spring 2020 or spring 2021. Benchmark assessment scores are converted to an estimated M-STEP proficiency category based on a linking studies from NWEA and Curriculum Associates. Proficiency rates from the 2018-19 and 2021-22 M-STEP include all students in districts that use the MAP Growth or i-Ready assessments.

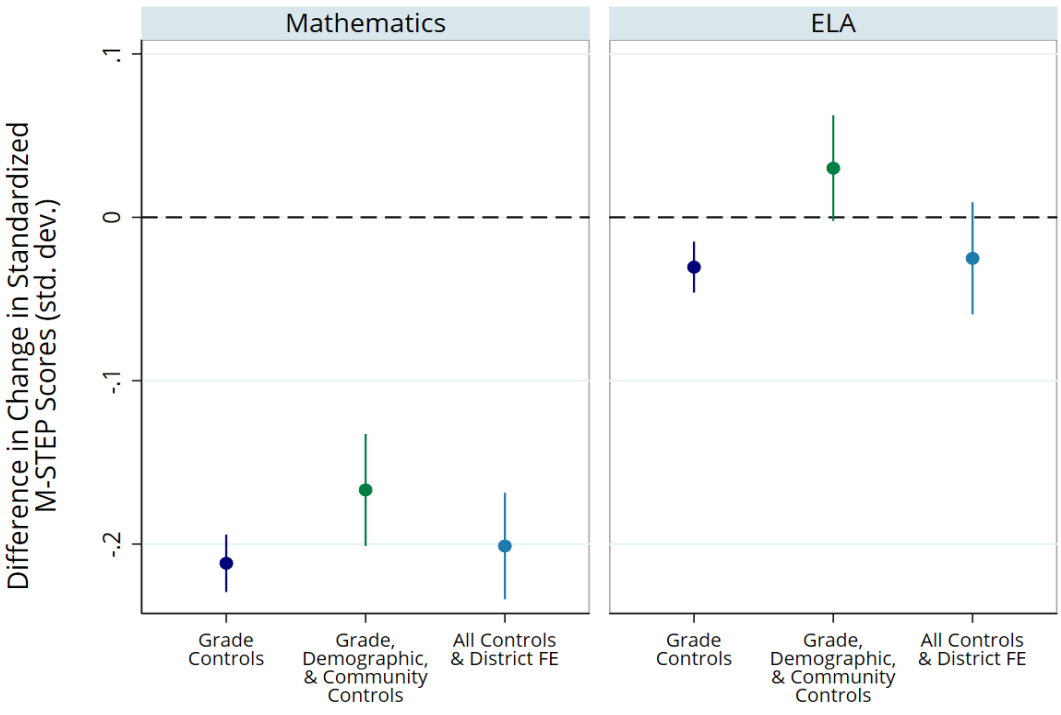
Figure 1 shows how the estimated distribution of M-STEP proficiency levels using benchmarks outcomes for Michigan 3rd- through 7th-grade students compares to the actual distribution of M-STEP outcomes from the spring 2019 and spring 2022 administrations of the assessment. As seen in the figure, achievement declined in the first pandemic year and remained lower than pre-pandemic levels in the next two years. Specifically, compared to students in the same districts who took the M-STEP in 2018-19, more students were classified as "not proficient" and fewer were classified as "proficient" or "advanced" based on their spring 2021 or spring 2022 benchmark assessment scores across both subjects. The percentages of students in each of these proficiency levels, however, did not change much between 2020-21 and 2021-22.

Importantly for our study, the estimated M-STEP proficiency rates from spring 2022 generally align with actual outcomes from the spring 2022 M-STEP administration. Indeed, the underlying correlation between individual benchmark and M-STEP scores from the spring 2022 administrations of both tests is 0.902 in math and 0.834 in reading/ELA. This suggests that we gain a similar signal from both measures of performance.

M-STEP Achievement Growth

Figures 2 through 5 provide our results from estimating model (1), examining differences in achievement growth between students in the pre-pandemic and pandemic M-STEP cohorts. Tables A.1.1 through A.1.3 provide the coefficient estimates from these models. In Figure 2, the zero-line represents the average three-year M-STEP growth for students in the pre-pandemic cohort, and we show results from models that initially control for students' grade level then sequentially add demographic/community characteristics and district fixed effects. In Figures 3 through 5, the zero-line represents the average three-year M-STEP growth for pre-pandemic cohort students in the specific reference group (i.e., White or non-economically disadvantaged students). Given the similarities between models that do and do not include district fixed effects, these latter figures only provide results from our preferred models that include district fixed effects.

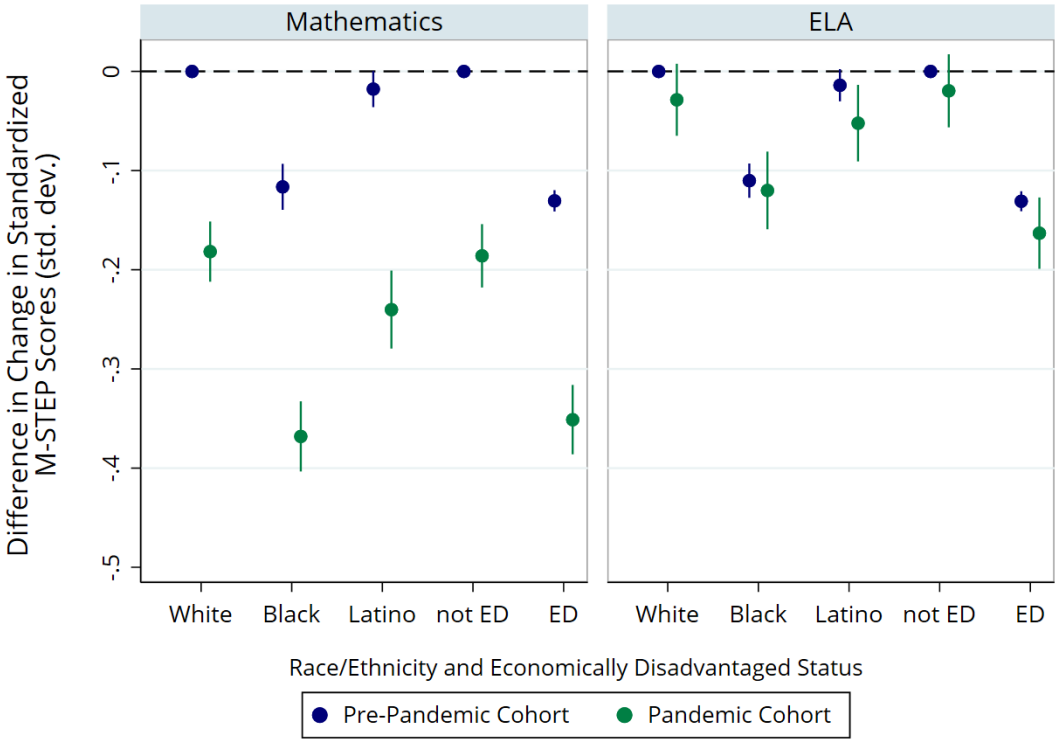
Figure 2. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



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 2 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 sd less in math over the three pandemic-affected years than did students in the pre-pandemic cohort. ELA growth for students in the pandemic cohort was generally similar to those 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 3. 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

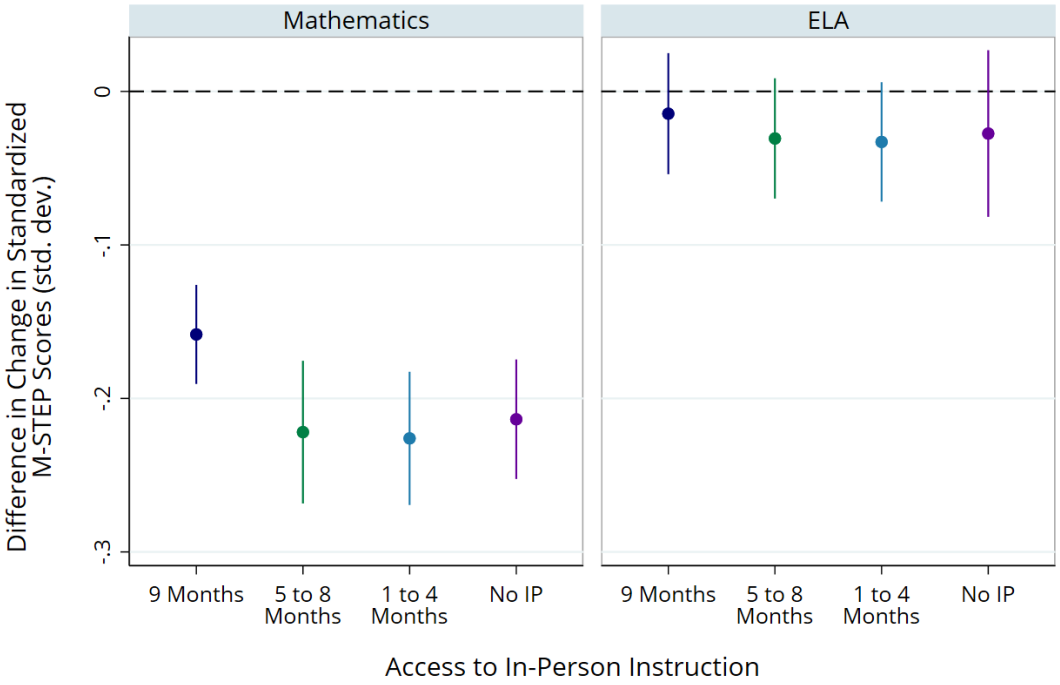


Notes: Each model includes student demographics and community characteristics, grade-level indicators for each sub-cohort to control for differences in learning trajectories between younger and older students, and district fixed effects to control for time-invariant, unobservable characteristics of each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

Figure 3 provides results from the district fixed effects model, this time examining heterogeneity by race/ethnicity and economically disadvantaged status. Even prior to the pandemic, disparities in achievement growth existed such that Black and 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 0.112 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 (-0.368 and -0.240 sd, respectively). Similarly, math achievement growth for economically disadvantaged students in the pre-pandemic cohort was 0.130 sd behind 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 for students in the pandemic cohort.

Figure 4 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.[5] 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 (-0.147 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 nearly 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 (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.

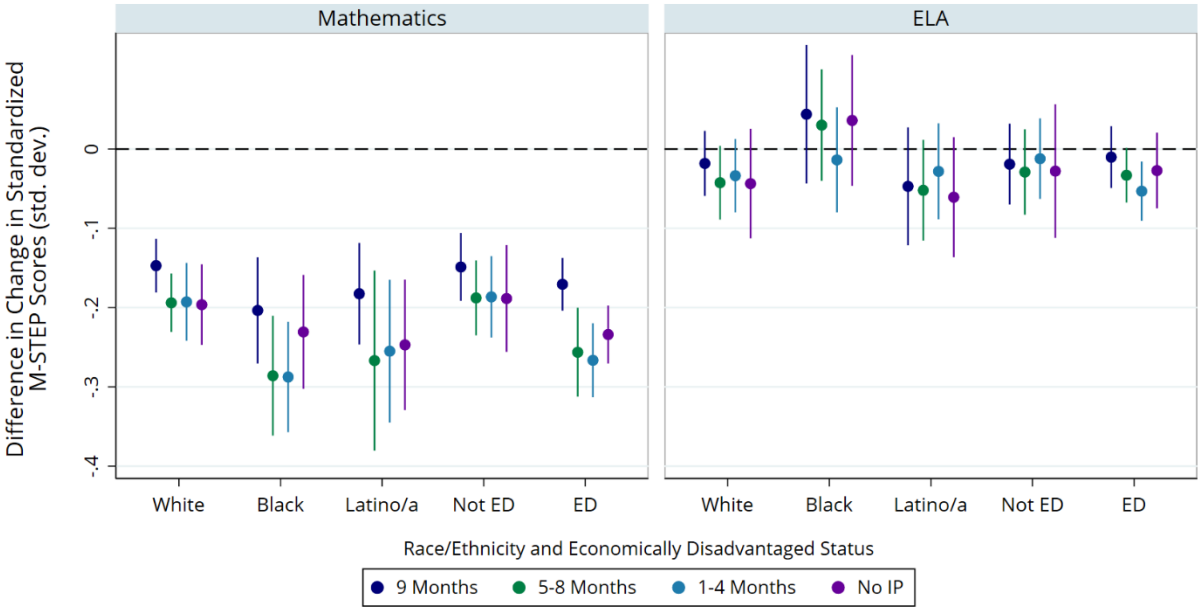
Figure 4. Differences in Learning Trajectories between Pre-Pandemic and Pandemic M-STEP Cohorts by 2020-21 Instructional Modality, 2016-2019 and 2019-2022 M-STEP Mathematics and ELA Assessments



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.

Finally, we find that learning remotely adversely affected all students regardless of their race/ethnicity or socioeconomic status. Figure 5 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 3. 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 relative to non-disadvantaged.

Figure 5. 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.

Benchmark Achievement Trends

Figures 6 through 9 show adjusted trends in standardized math and reading benchmark achievement for students who started the 2020-21 school year in grades five through seven and completed a MAP Growth or i-Ready assessment in all four administration periods during the 2020-21 and 2021-22 school years. 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.[6] Since the trends in each specification are generally similar, we only report estimates for models that include district fixed effects in Figures 6 through 9 and Tables A.1.5 and A.1.6.

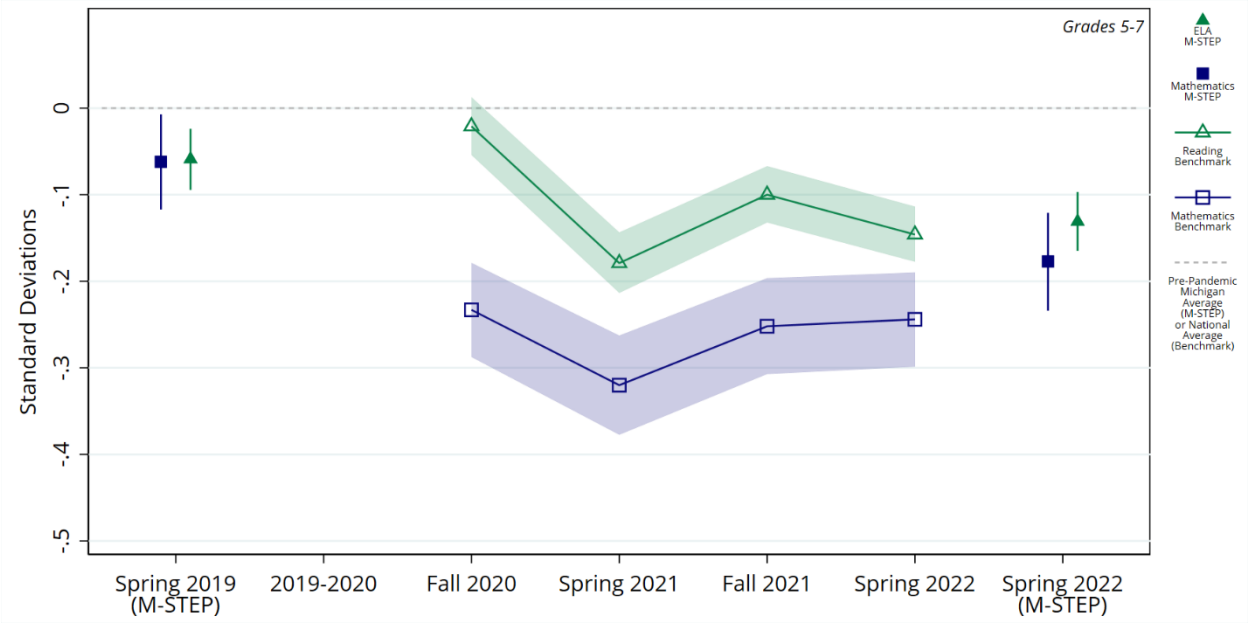
As noted earlier, benchmark assessment scores are standardized relative to pre-pandemic national norms for each grade, subject, and testing period. As such, we interpret the trend lines in Figures 6 through 9 as deviations from the average scores for nationally representative samples of students who 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 6. First, by fall 2020, average benchmark scores in Michigan were below the pre-pandemic norms for both reading and math (0.021 and 0.233 sd below average, respectively). Again, although the spring 2019 M-STEP and fall 2020 data points are not directly comparable, it is clear that Michigan students in our sample were performing only slightly better in reading in fall 2020 than in spring 2019 but were substantially behind in math. Second, we find that both math and reading benchmark scores dropped considerably during the 2020-21 school year, falling even farther behind the national pre-pandemic norm. Between spring and fall 2021, however, Michigan students experienced faster than expected growth, such that by fall 2021 they had almost caught up to where they had started the 2020-21 school year in math, but still trailed their fall 2020 average score in reading. Nonetheless, these scores both remained substantially behind the average standardized M-STEP score from spring 2019. Then, during the 2021-22 school year, students made slightly higher than expected progress in math relative to the pre-pandemic national norm, whereas reading achievement fell relative to the national norm once again, albeit at a much slower rate than the prior year.[7] Spring 2022 benchmark and M-STEP scores were generally similar in both subjects.

Thus, overall, total 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 exams 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. Worrisomely, however, there is an indication that recovery stalled in the 2021-22 school year, as a continued upward trend between fall 2021 and spring 2022 would be necessary to recover all of the losses from the early part of the pandemic. It is unclear at this time whether the recovery has accelerated into 2022-23.

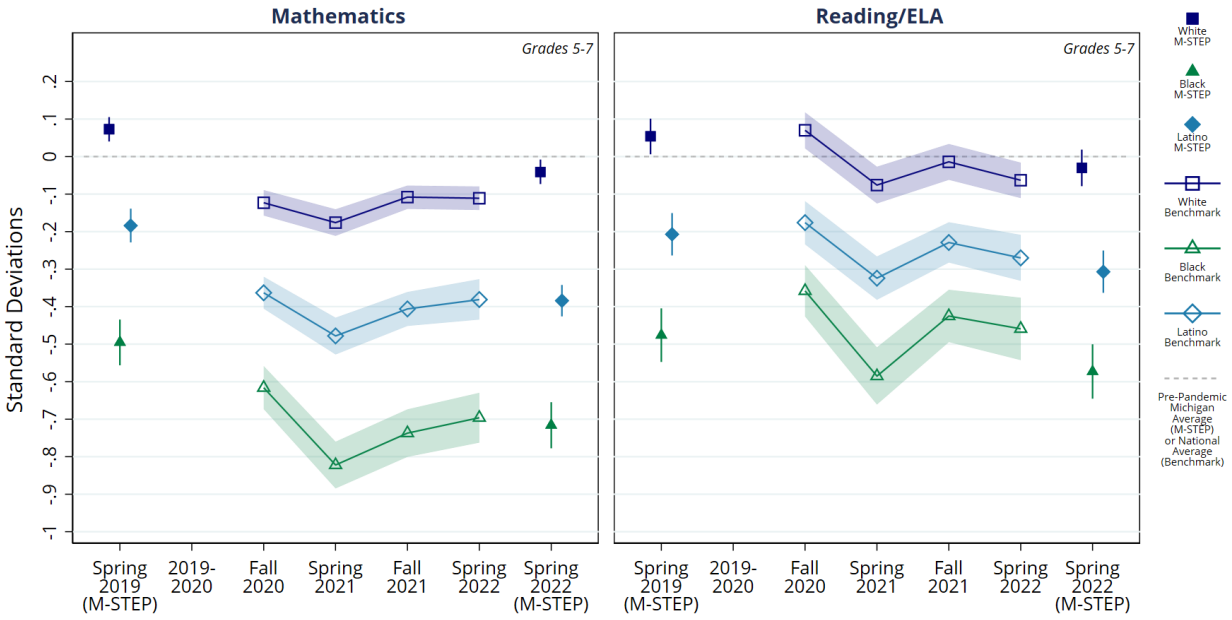
Figure 6. Regression Adjusted Scale Score Trends, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-7



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 and 2022 M-STEP estimates have been standardized relative to national norms.

Figure 7 shows differences in adjusted trends in standardized math and reading benchmark achievement by race/ethnicity. We find similar patterns across subgroups, all in line with the overall results shown in Figure 6. White students had consistently higher scores in both subjects compared to Black and Latino students, with Black students scoring the lowest of the three subgroups. 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. The declines in 2020-21 were largest for Black students (-0.207 and -0.227 sd in math and reading, respectively), followed by Latino students (-0.116 and -0.148 sd), and White students (-0.053 and -0.146 sd). During the 2021-22 school year, these gaps in math began to diminish, as math achievement for White students plateaued whereas Black and Latino math achievement increased slightly by 0.041 and 0.026 sd, respectively. Reading achievement decreased across all three groups of students during the 2021-22 school year (between approximately -0.035 and -0.050 sd across all student groups).

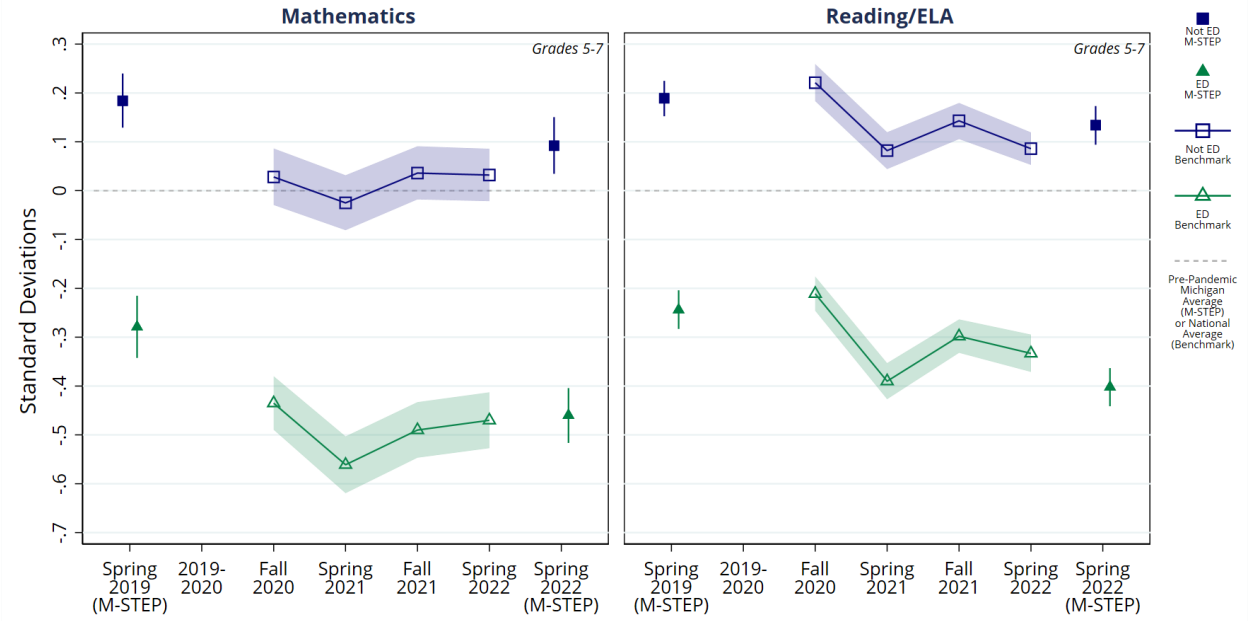
Figure 7. Regression Adjusted Scale Score Trends by Race/Ethnicity, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-7



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 and 2022 M-STEP estimates have been standardized relative to national norms.

These patterns reveal that the pandemic exacerbated racial/ethnic math achievement gaps. In fall 2020, the differences in math achievement between White students and their Black and Latino peers were 0.492 and 0.240 sd, respectively. By spring 2022, the White-Black and White-Latino gaps increased to 0.585 and 0.270 sd, respectively. In reading, the White-Black and White Latino gaps both decreased slightly by approximately 0.04 sd between fall 2020 and spring 2022.

Figure 8. Regression Adjusted Scale Score Trends by Economically Disadvantaged Status, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-7



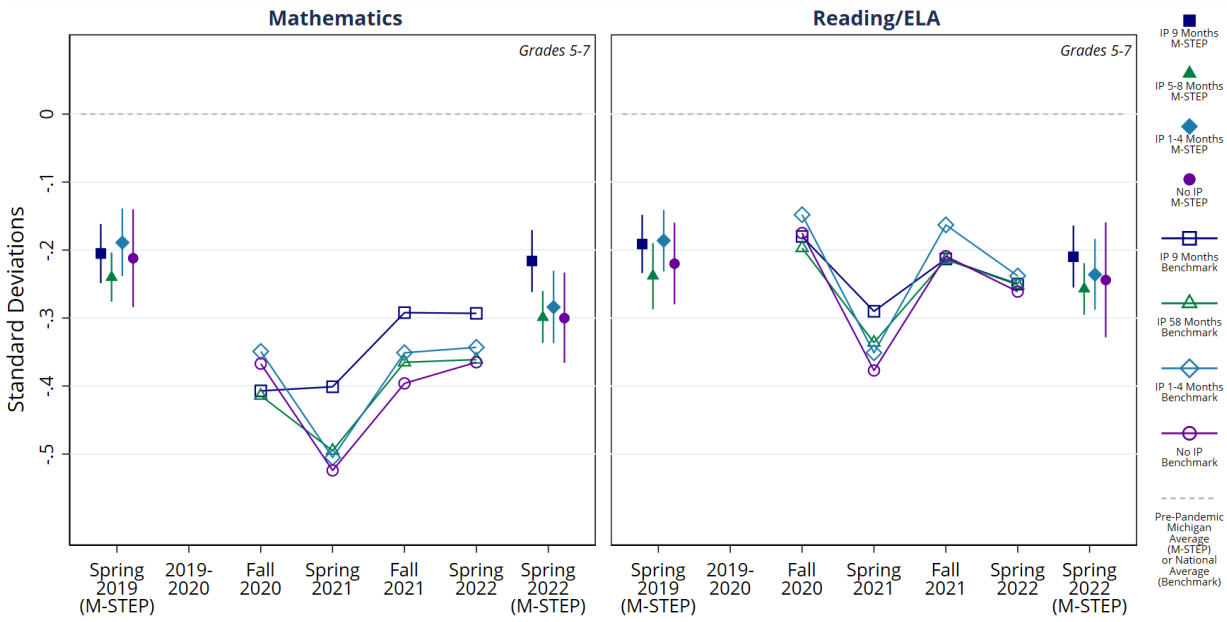
Notes: These regression estimates include only students with benchmark assessment scores for every possible testing period. Each model controls for student demographics. Test scores have been standardized relative to NWEA's and Curriculum Associates' pre-pandemic national norms. Spring 2019 and 2022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

Figure 8 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, followed by a rebound in scores during the 2021-22 school year. The decreases in 2020-21 for both subjects were slightly larger for economically disadvantaged students (-0.126 to -0.179 sd in math and reading, respectively) than more advantaged students (-0.053 and -0.139 sd). Math achievement for the more advantaged students plateaued during the 2021-22 school year, while economically disadvantaged math achievement increased by 0.019 sd. Reading achievement across both groups students decreased during the 2021-22 school year (-0.035 and -0.056 sd for economically disadvantaged and non-economically disadvantaged students, respectively).

Similar to results for disparities by race and ethnicity, we find that the pandemic exacerbated math achievement gaps by economically disadvantaged status. In fall 2020, economically disadvantaged students scored 0.463 sd below their peers in

math. By spring 2022, this gap increased to 0.502 sd. The same gap in reading decreased slightly, from 0.432 to 0.419 sd between fall 2020 and spring 2022.

Figure 9. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-7



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 and 2022 M-STEP estimates have been standardized relative to national norms.

Finally, Figure 9 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 Figure 9 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 9 are generally not statistically significant. We provide an additional figure in the Appendix where the confidence intervals are included (Figure A.1.3).

As might be expected given the overlapping confidence intervals, we find few differences in achievement across districts that offered varying levels of in-person instruction conditional on having hybrid or remote modalities during 2020-21, but students in districts that offered in-person instruction throughout all of 2020-21 performed better during that school year, consistent with our earlier results from analyses of the M-STEP scores. Initially, districts that offered in-person instruction

throughout all of the 2020-21 school year had close to the lowest achievement levels as of the fall 2020 benchmarks, though these differences were not statistically significant. However, while schools that were in-person all year maintained math growth equivalent to pre-pandemic national norms between fall 2020 and spring 2021, those with remote schooling for any part of the year saw large declines in achievement relative to these norms, regardless of the number of months in which remote modality was offered. After the 2020-21 school year, as schools returned to mostly in-person learning, math achievement growth equalized across modalities. Thus, throughout the first two pandemic years, schools that remained entirely in person saw smaller overall interruptions to math learning relative to the 2019-20 M-STEP scores, consistent with the pre- vs. post-pandemic M-STEP comparisons in Figure 4. For ELA, initially between Fall 2020 and spring 2021, districts that were entirely in-person performed better than those in other modalities, but then districts with remote instruction caught up such that there was little difference by modality in ELA by spring 2022, again consistent with the results in Figure 4.

DISCUSSION

Our M-STEP results suggest that, while middle-school 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 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. We find that, early in the pandemic, Michigan student achievement on benchmark assessments was already below national norms. In the first full pandemic-impacted school year (2020-21), achievement trends for Michigan middle school students fell further behind national norms before partially rebounding during the 2021-22 school year, especially for math. However, although math achievement growth began to mirror pre-pandemic trends in the 2021-22 school year, this is insufficient to enable students to "catch up" to where they would have been prior to the pandemic. 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. Whether or not we see this accelerated growth will become apparent as 2022-23 academic year data become available.

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, grew between fall 2020 and spring 2022. 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 achievement in the 2021-22 school year. There was no discernable difference in reading student achievement by spring of 2022.

Together, these summative and formative assessment results paint a nuanced picture of student achievement trends and outcomes during the full school years most impacted by the COVID-19 pandemic. By the spring of 2022, we see persistent and large negative effects on math. Effects on ELA were generally small and statistically insignificant. Benchmark results make clear that during the initial phase of the pandemic, when school buildings first shuttered for in-person learning across the state, Michigan students were scoring much further below national norms in math than they were in reading. While both math and reading achievement were negatively impacted in the 2020-21 school year, students in our data improved at a rate higher than national norms would have predicted over the summer of 2021. During the 2021-22 school year, math achievement improved relative to national norms and students were able to recoup achievement losses in 2020-21. However, given how far below national norms Michigan students' math scores had fallen by fall 2020, even these relative improvements in the second full year of the pandemic were insufficient to allow students to rebound completely.

We make several recommendations for policymakers and educators based on these findings. First, results from the 2021-22 school year make clear that the road to academic recovery in Michigan will not be quick and a return to "business as normal" will be insufficient to improve student achievement to pre-pandemic levels. For example, based on benchmark outcomes from spring 2022 and typical growth measures defined by each assessment provider, 5th- through 7th-grade students in Michigan will need to achieve roughly 140 to 180% of typical fall-to-spring growth in math, and between 120 to 140% of typical growth in reading, during the 2022-23 school year to reach the 50th percentile of pre-pandemic achievement by spring 2023. Clearly, then, the tremendous effect that the COVID-19 pandemic has had, and continues to have, on student learning will not be addressed quickly or without a substantial and sustained influx of resources to support education in Michigan.

These patterns in achievement and achievement growth mirror recent findings from across the U.S. (e.g., Curriculum Associates, 2022; Goldhaber et al., 2022; Kuhfeld & Lewis, 2022). It will be critical for local, state, and the federal governments to prioritize both short- and longer-term investments into public education as educators and students work to recover from the trauma of the COVID-19 pandemic.

Moreover, our and others' results show particularly troublesome disruptions to math achievement. However, there has been relatively little discussion of ways to improve math achievement (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.

Thus, 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 allows policymakers, educators, and stakeholders to assess progress in the coming years. In particular, research exploring trends in academic achievement over the past two 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 between 2020-21 and 2021-22, disparities between each group persist. Any decisions to reduce monitoring of student learning progress may exacerbate longstanding achievement gaps.

In sum, our results bolster other data from around the country that make clear the road to recovery 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 in order 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, in Michigan and elsewhere.

END NOTES

[1] Most districts provided authorization for us to construct district-level aggregate datasets from their student-level benchmark assessment data, while some districts chose to only provide aggregate datasets they prepared themselves. Districts that chose to aggregate their own data were instructed to calculate average scale scores across all students in the same subgroup and grade level who completed an assessment from the same provider in each of the four testing periods. We apply the same sample restrictions and construct equivalent aggregate measures for the districts that provided student-level data, then compile all districts' aggregate data into a combined dataset for the benchmark analysis.

[2] 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.

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

[4] 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 access.

[5] 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.

[6] Estimates without district fixed effects are similar and available by request.

[7] To better see why a flat line indicates "normal" growth, Appendix Figures A.1.1 and A.1.2 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.

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

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/education/our-insights/teacher-survey-learning-loss-is-global-and-significant

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.

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. https://www.curriculumassociates.com/-/media/mainsite/files/corporate/state-of-student-learning-2022.pdf?_ga=2.249214663.1791340512.1663161051-741382319.1663161051

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, 23328584221128035.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2021, July 27). COVID-19 and education: The lingering effects of unfinished learning. McKinsey & Company. https://www.mckinsey.com/industries/education/our-insights/covid-19-and-education-the-lingering-effects-of-unfinished-learning

Fensterwald, J. (2020, November 30). Early data on learning loss show big drop in math, but not reading skills. EdSource. https://edsource.org/2020/early-dataon-learning-loss-show-big-drop-in-math-but-not-reading-skills/644416

Ferren, M. (2021, July 6). Remote learning and school reopenings: What worked and what didn't. Center for American Progress. https://www.americanprogress.org/issues/education-k12/reports/2021/07/06/501221/remote-learning-school-reopenings-worked-didnt/

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. https://doi.org/10.1007/s11528-021-00625-5

Goldhaber, D., Kane, T. J., McEachin, A., Morton, E., Patterson, T., & Staiger, D. O. (2022). The consequences of remote and hybrid instruction during the pandemic (No. w30010). National Bureau of Economic Research.

Greater Fort Lauderdale Alliance. (2022, June 29). State releases 2022 Florida standards assessment and end of course exam results. The Alliance. https://www.gflalliance.org/news/2022/06/29/education-news/state-releases2022-florida-standards-assessment-and-end-of-course-examresults/

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. https://www.rand.org/pubs/research_reports/RRA168-2.html

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. https://epicedpolicy.org/fall-2020-covid-19-survey_policy_brief/

Idaho State Department of Education. (2022, July 6). Early reading test shows statewide gains from previous spring and fall. [Press release]. https://www.sde.idaho.gov/communications/files/news-releases/07-06-2022-Idaho-early-reading-test-shows-strong-statewide-gains-from-previous-springand-fall.pdf

Jack, R., Halloran, C., Okun, J., & Oster, E. (2022). Pandemic schooling mode and student test scores: evidence from US school districts. American Economic Review: Insights.

Kilbride, T., Hopkins, B., Strunk, K.O. & Yu, D. (2022) Michigan's 2020-21 and 2021-22 Benchmark Assessments. Education Policy Innovation Collaborative. https://epicedpolicy.org/wp-content/uploads/2022/10/COVID_Benchmark_Assessments_Report_Oct2022.pdf.

Kogan, V. (2022, June). Academic achievement and pandemic recovery in Ohio: An update from fall third grade English language arts assessments. https://glenn.osu.edu/academic-achievement-pandemic-recovery

Kogan, V., & Lavertu, S. (2021, January 27). The COVID-19 pandemic and student achievement on Ohio's third-grade English language arts assessment. John Glenn College of Public Affairs, Ohio State University. http://glenn.osu.edu/educational-governance/reports/reportsattributes/ODE_ThirdGradeELA_KL_1-27-2021.pdf

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. https://www.brookings.edu/blog/brown-center-chalkboard/2022/03/03/the-pandemic-has-had-devastating-impacts-on-learning-what-will-it-take-to-help-students-catch-up/

Levin, K. (2021, March 18). Michigan lost 62,000 students this fall. Black enrollment fell 5%. Detroit Free Press. https://www.freep.com/story/news/education/2021/03/17/michigan-publicschools-enrollment-decline/4730513001/

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. https://www.the74million.org/article/public-school-enrollment-down3-percent-worst-century/

Michigan Department of Education (2019). *Technical Report, Spring 2019 Michigan Student Test of Educational Progress (M-STEP).* https://www.michigan.gov/mde/-/media/Project/Websites/mde/Year/2020/05/14/Spring_2019_M-STEP_Technical_Report_Main_report.pdf?rev=0c3bbf9629ea411395157233c92439e0&hash=7889BF8A50495DC27B1FD1AD62820C04

Michigan Public Act 147 of 2020, Mich., MCL § 388.1621f (2020) http://www.legislature.mi.gov/(S(rhqyboiv2ivksh4xj0wbiqpl))/mileg.aspx?page=GetObject&objectname=mcl-388-1621f

Michigan Public Act 148 of 2020, Mich. MCL § 388.1701 (2020) http://www.legislature.mi.gov/(S(hotihihl5cde5dkkjzjvx4g3))/mileg.aspx?page=getObject&objectName=mcl-388-1701

Michigan Public Act 149 of 2020, Mich., MCL § 388.1606 (2020) http://legislature.mi.gov/doc.aspx?mcl-388-1606

National Center for Education Statistics (2022). NAEP long-term trend assessment results: Reading and mathematics. The Nation's Report Card. https://www.nationsreportcard.gov/highlights/ltt/2022/

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. https://www.nwea.org/content/uploads/2016/12/MI-MAP-Growth-LinkingStudy-Report_NWEA_2020-12-22.pdf

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

https://www.edweek.org/leadership/more-than-1-million-students-didntenroll-during-the-pandemic-will-they-come-back/2021/06

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. https://edpolicyinca.org/newsroom/covid-19-and-educational-equity-crisis

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. https://www.air.org/sites/default/files/2021-08/research-brief-covid-surveypersistent-challenges-july-2021rev.pdf

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. https://www.edweek.org/technology/extreme-chronic-absenteeismpandemic-school-attendance-data-is-bleak-but-incomplete/2021/07

Smarter Balanced Assessment Consortium (2020). *2018-19 summative technical report.* https://technicalreports.smarterbalanced.org/2018-19 summative-report/_book/

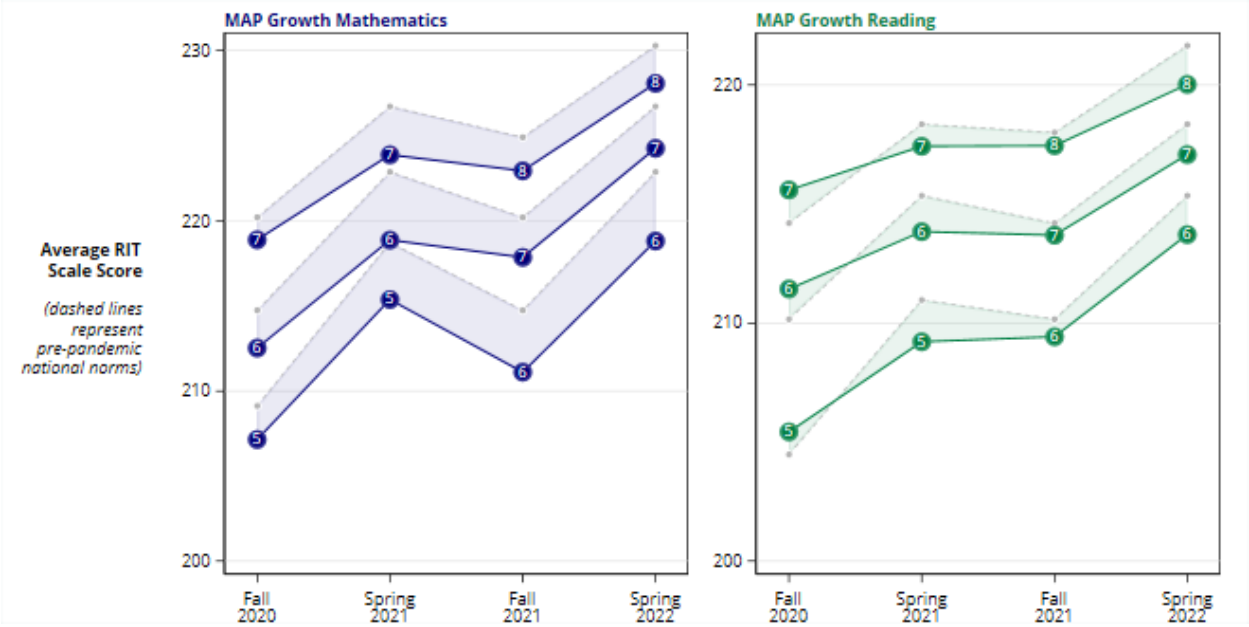
Tennessee Department of Education. (2022, June 14). Tennessee Releases 2021–22 TCAP State-Level Results Highlighting Significant Learning Acceleration. https://www.tn.gov/education/news/2022/6/14/tennessee-releases-2021-22-tcap-state-level-results-highlighting-significant-learning-acceleration-.html

Texas Education Agency. (2022, July 1). TEA Releases 2022 Grades 3–8 STAAR Results | Texas Education Agency. https://tea.texas.gov/about-tea/news-andmultimedia/news-releases/news-2022/tea-releases-2022-grades-3-8-staarresults

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. https://www.crpe.org/publications/how-much-havestudents-missed-academically-because-pandemic-review-evidence-date

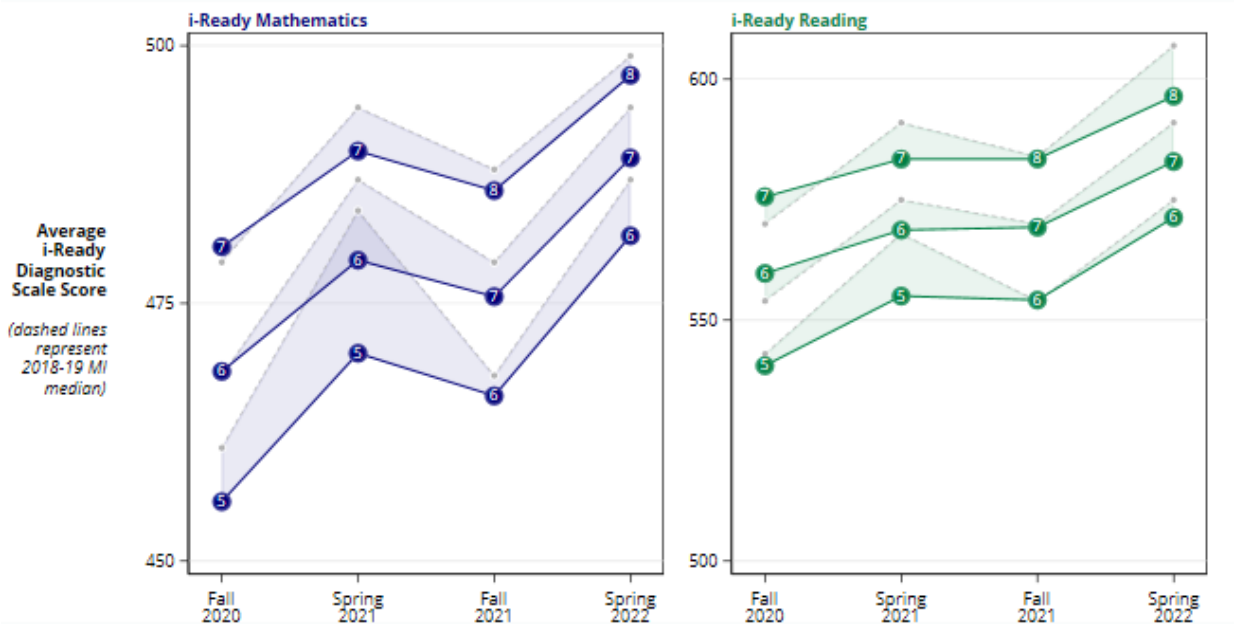
APPENDIX A.1. SUPPLEMENTARY FIGURES AND TABLES

Figure A.1.1. Trends in Average Scale Scores, NWEA MAP Growth, Grades 5-7, Fall 2020 to Spring 2022



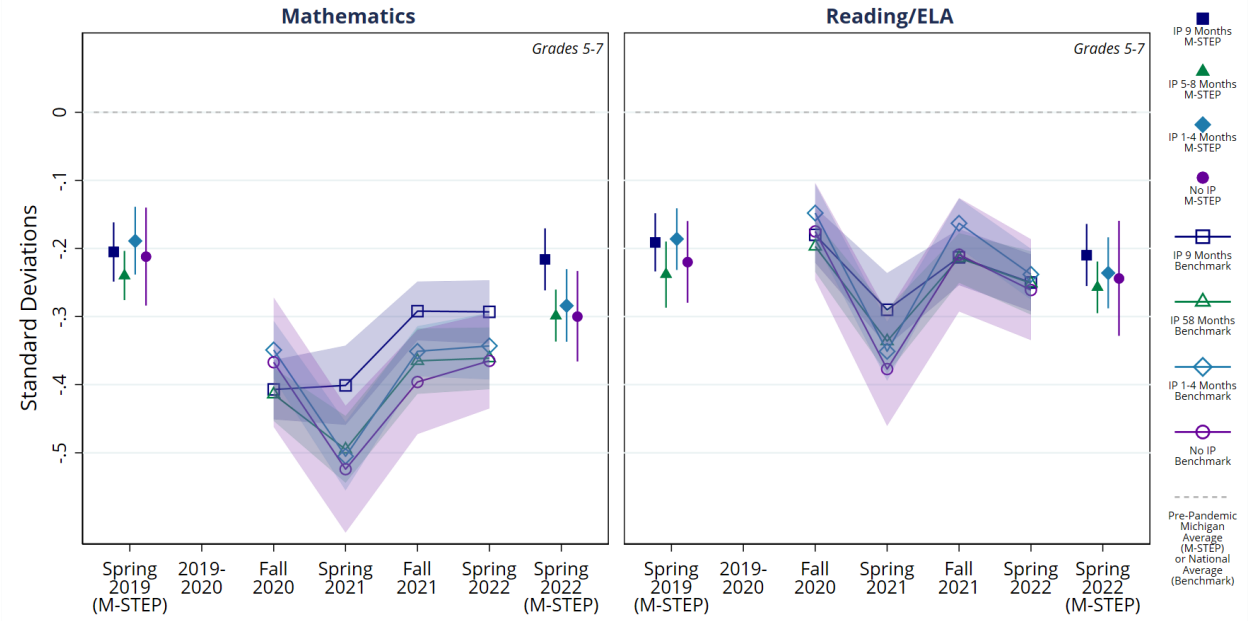
Notes: These averages include only students with benchmark assessment scores for every possible testing period. The comparison points in the figure represent the 50th percentile of NWEA's conditional growth distribution. RIT stands for Rasch unit scale.

Figure A.1.2. 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.1.3. Regression Adjusted Scale Score Trends by 2020-21 Instructional Modality, NWEA MAP Growth and Curriculum Associates' i-Ready, Grades 5-7



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 and 2022 M-STEP estimates have been standardized relative to national norms.

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***	-0.167***	-0.201***	-0.030***	0.030+	-0.025
	(0.009)	(0.017)	(0.017)	(0.008)	(0.016)	(0.017)
Black		-0.188***	-0.150***		-0.071***	-0.101***
		(0.010)	(0.007)		(0.010)	(0.007)
Latino		-0.069***	-0.037***		-0.034**	-0.019*
		(0.010)	(0.007)		(0.011)	(0.008)
Economically Disadvantaged		-0.199***	-0.147***		-0.176***	-0.137***
		(0.006)	(0.004)		(0.007)	(0.004)
Base-Year Achievement	-0.158***	-0.241***	-0.247***	-0.221***	-0.280***	-0.285***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)	(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

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 (0.018)	-0.020 (0.019)	-0.014 (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

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***	-0.204***	-0.183***	-0.149***	-0.171***	-0.018	0.044	-0.047	-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**	-0.082*	-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

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 5-7						
	Mathematics			Reading		
	(1)	(2)	(3)	(4)	(5)	(6)
Spring 2019 (M-STEP)	0.165***	0.179***	0.171***	-0.046***	-0.033**	-0.038***
	(0.010)	(0.012)	(0.010)	(0.009)	(0.011)	(0.010)
Spring 2021	-0.082***	-0.082***	-0.087***	-0.154***	-0.151***	-0.158***
	(0.009)	(0.013)	(0.012)	(0.006)	(0.011)	(0.008)
Fall 2021	-0.045***	0.048***	-0.019	-0.106***	-0.026*	-0.079***
	(0.009)	(0.013)	(0.013)	(0.005)	(0.010)	(0.008)
Spring 2022	-0.038***	0.055***	-0.011	-0.152***	-0.073***	-0.125***
	(0.011)	(0.014)	(0.012)	(0.011)	(0.014)	(0.010)
Spring 2022 (M-STEP)	0.029**	0.122***	0.056***	-0.137***	-0.058***	-0.110***
	(0.010)	(0.014)	(0.012)	(0.008)	(0.012)	(0.011)
i-Ready	-0.205	-0.032	0.073	-0.198	-0.046+	-0.005
	(0.214)	(0.024)	(0.133)	(0.194)	(0.025)	(0.080)
Black, District Percent		-0.439***	-0.506**		-0.291***	-0.247+
		(0.041)	(0.173)		(0.044)	(0.150)
Latino, District Percent		0.119	0.016		0.220*	0.186
		(0.127)	(0.171)		(0.112)	(0.144)
ED, District Percent		-1.065***	-0.145*		-0.976***	-0.082
		(0.053)	(0.072)		(0.048)	(0.067)
Constant	-0.116**	-0.395***	-0.233***	0.0676*	-0.176***	-0.021
	(0.035)	(0.018)	(0.028)	(0.0290)	(0.016)	(0.017)
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	0.054	0.790	0.906	0.051	0.759	0.888

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 and 2022 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 5-7				
	Math		Reading	
	(1)	(2)	(3)	(4)
Spring 2019 (M-STEP)	0.196*** (0.010)	0.156*** (0.014)	-0.016+ (0.009)	-0.033** (0.011)
Spring 2021	-0.053*** (0.012)	-0.053*** (0.016)	-0.146*** (0.008)	-0.139*** (0.009)
Fall 2021	0.015 (0.011)	0.008 (0.013)	-0.084*** (0.008)	-0.079*** (0.008)
Spring 2022	0.012 (0.013)	0.004 (0.016)	-0.133*** (0.010)	-0.135*** (0.009)
Spring 2022 (M-STEP)	0.083*** (0.013)	0.064*** (0.015)	-0.100*** (0.010)	-0.088*** (0.011)
Black*Spring 2019 (M-STEP)	-0.568*** (0.033)		-0.530*** (0.034)	
Black*Fall 2020	-0.492*** (0.033)		-0.428*** (0.034)	
Black*Spring 2021	-0.646*** (0.033)		-0.509*** (0.039)	
Black*Fall 2021	-0.629*** (0.033)		-0.410*** (0.035)	
Black*Spring 2022	-0.585*** (0.038)		-0.396*** (0.045)	
Black*Spring 2022 (M-STEP)	-0.605*** (0.033)		-0.509*** (0.037)	
Latino*Spring 2019 (M-STEP)	-0.257*** (0.023)		-0.261*** (0.022)	
Latino*Fall 2020	-0.240*** (0.021)		-0.246*** (0.023)	
Latino*Spring 2021	-0.303*** (0.023)		-0.248*** (0.023)	
Latino*Fall 2021	-0.298*** (0.022)		-0.215*** (0.022)	
Latino*Spring 2022	-0.270*** (0.029)		-0.206*** (0.029)	
Latino*Spring 2022 (M-STEP)	-0.273*** (0.021)		-0.243*** (0.023)	
ED*Spring 2019 (M-STEP)		-0.433*** (0.018)		-0.445*** (0.016)
ED*Fall 2020		-0.463*** (0.022)		-0.432*** (0.017)
ED*Spring 2021		-0.536*** (0.017)		-0.472*** (0.017)
ED*Fall 2021		-0.526*** (0.017)		-0.440*** (0.017)
ED*Spring 2022		-0.502*** (0.020)		-0.419*** (0.019)

ED*Spring 2022 (M-STEP)		-0.492*** (0.018)		-0.488*** (0.019)
i-Ready	-0.0285 (0.0541)	0.068 (0.127)	-0.048 (0.122)	-0.006 (0.077)
Constant	-0.1232*** (0.0173)	0.029 (0.029)	0.070** (0.024)	0.221*** (0.019)
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	0.861	0.885	0.823	0.864

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 and 2022 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 5-7		
	Math	**Reading**
Spring 2019 (M-STEP)	0.202***	-0.011
	(0.022)	(0.018)
Spring 2021	0.007	-0.109***
	(0.024)	(0.014)
Fall 2021	0.116***	-0.032**
	(0.014)	(0.012)
Spring 2022	0.114***	-0.070***
	(0.022)	(0.014)
Spring 2012 (M-STEP)	0.191***	-0.029+
	(0.018)	(0.016)
IP 5-8 Months*Spring 2019 (M-STEP)	-0.035	-0.047+
	(0.026)	(0.028)
IP 5-8 Months*Fall 2020	-0.007	-0.016
	(0.023)	(0.021)
IP 5-8 Months*Spring 2021	-0.094***	-0.047+
	(0.027)	(0.025)
IP 5-8 Months*Fall 2021	-0.074*	-0.001
	(0.029)	(0.024)
IP 5-8 Months*Spring 2022	-0.068*	-0.000
	(0.031)	(0.030)
IP 5-8 Months*Spring 2022 (M-STEP)	-0.083**	-0.048+
	(0.027)	(0.026)
IP 1-4 Months*Spring 2019 (M-STEP)	0.017	0.005
	(0.031)	(0.028)
IP 1-4 Months*Fall 2020	0.059*	0.033
	(0.025)	(0.024)
IP 1-4 Months*Spring 2021	-0.104***	-0.061*
	(0.028)	(0.025)
IP 1-4 Months*Fall 2021	-0.059*	0.050*
	(0.025)	(0.024)
IP 1-4 Months*Spring 2022	-0.050	0.013
	(0.032)	(0.025)
IP 1-4 Months*Spring 2022 (M-STEP)	-0.068*	-0.026
	(0.034)	(0.033)
IP 0 Months*Spring 2019	-0.007	-0.029
	(0.043)	(0.037)
IP 0 Months*Fall 2020	0.040	0.006
	(0.051)	(0.039)
IP 0 Months*Spring 2021	-0.123**	-0.088*
	(0.046)	(0.042)
IP 0 Months*Fall 2021	-0.105*	0.003
	(0.042)	(0.045)
IP 0 Months*Spring 2022	-0.072+	-0.010
	(0.041)	(0.042)

IP 0 Months*Spring 2022 (M-STEP)	-0.084*	-0.034
	(0.039)	(0.047)
i-Ready	-0.022	-0.038
	(0.025)	(0.025)
Constant	-0.407***	-0.180***
	(0.022)	(0.021)
District-Level Student Controls	Y	Y
Grade Controls	Y	Y
COVID-19 Death Rates	Y	Y
District Fixed Effects	Y	Y
R^2	0.799	0.766

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 and 2022 M-STEP estimates have been standardized relative to national norms. $^{+}$ *p < 0.10,* * *p < 0.05,* ** *p < 0.01,* *** *p < 0.001*